

# An approximate inverse to the extended Born modeling operator

Jie Hou<sup>1</sup> and William W. Symes<sup>1</sup>

## ABSTRACT

Given a correct (data-consistent) velocity model, reverse time migration (RTM) correctly positions reflectors but generally with incorrect amplitudes and wavelets. Iterative least-squares migration (LSM) corrects the amplitude and wavelet by fitting data in the sense of Born modeling, that is, replacing migration by Born inversion. However, LSM also requires a correct velocity model, and it may require many migration/demigration cycles. We modified RTM in the subsurface offset domain to create an asymptotic (high-frequency) approximation to extended LSM. This extended Born inversion operator outputs extended reflectors (depending on the subsurface offset and position in the earth) with correct amplitude and phase, in the sense that similarly extended Born modeling reproduces the data to good accuracy. Although

the theoretical justification of the inversion property relies on ray tracing and stationary phase, application of the weight operators does not require any computational ray tracing. The computational expense of the extended Born inversion operator is roughly the same as that of extended RTM, and the inversion (data-fit) property holds even when the velocity is substantially incorrect. The approximate inverse operator differs from extended RTM only in the application of data- and model-domain weight operators, and takes the form of an adjoint in the sense of weighted inner products in data and model space. Because the Born modeling operator is approximately unitary with respect to the weighted inner products, a weighted version of conjugate gradient iteration dramatically accelerates the convergence of extended LSM. An approximate LSM may be extracted from the approximate extended LSM by averaging over subsurface offset.

## INTRODUCTION

Reverse time migration (RTM) (Baysal et al., 1983; Loewenthal and Mufti, 1983; Whitmore, 1983) produces kinematically accurate short-scale reflectivity, with reflectors positioned as accurately as migration velocity permits. RTM images may be degraded by amplitude anomalies, low-frequency noise, and wavelet side lobes (Mulder and Plessix, 2004; Bednar and Bednar, 2006). Least-squares migration (LSM) (Nemeth et al., 1999; Dutta et al., 2014) compensates for all of these defects to some extent. LSM is actually a linearized inversion (Bourgeois et al., 1989), that is, a method for choosing short-scale reflectivity as a perturbation of the migration velocity model, so as to achieve a sample-by-sample mean-square best fit to data via Born modeling. It has been formulated as an iterative process involving repeated migrations and demigrations (Nemeth et al., 1999; Dutta et al., 2014), in an asymptotic approximation via generalized Radon transform inversion (Beylkin, 1985; Bleistein, 1987; De Hoop and Bleistein, 1997), and as a true amplitude modification of wave-equation migration, using one-way (Zhang et al., 2005) and two-way (Zhang et al., 2007) propagators.

In the following pages, we describe a modification of RTM that approximately inverts the space shift extension of Born forward modeling (Rickett and Sava, 2002; Symes, 2008). That is, it produces an image volume depending not just on position in the subsurface but also on a subsurface offset parameter (vector, in 3D). By analogy with the usual terminology, one might term this approximate linearized inversion as an extended LSM. Averaging the output over offset with a weight equal to one at zero offset produces an approximate inversion of the ordinary Born modeling operator, that is, an approximate LSM, provided that the velocity model is consistent with the data.

A concise mathematical description of this approximate inversion operator takes the form

$$\bar{F}^\dagger[v_0] = W_{\text{model}}[v_0]\bar{F}^*[v_0]W_{\text{data}}[v_0], \quad (1)$$

where  $\bar{F}^\dagger[v_0]$  is the approximate inverse of the extended Born modeling operator  $\bar{F}[v_0]$  at background velocity  $v_0$ ;  $\bar{F}^*[v_0]$  is the extended RTM operator, that is, the adjoint or transpose of the modeling operator  $\bar{F}[v_0]$ ;  $W_{\text{model}}[v_0]$  and  $W_{\text{data}}[v_0]$  are model- and data-domain

Manuscript received by the Editor 15 December 2014; revised manuscript received 2 July 2015; published online 2 September 2015.

<sup>1</sup>Rice University, The Rice Inversion Project, Houston, Texas, USA. E-mail: jie.hou@rice.edu; symes@caam.rice.edu.

© 2015 Society of Exploration Geophysicists. All rights reserved.

weighting operators, defined in detail in the “Theory” section. The central result of this paper is that an approximate inverse in the sense of high-frequency asymptotics takes this form, with weight operators having *explicitly computable* integral forms, depending only on phase space variables and the velocity field. The derivation of this remarkable fact is outlined in the “Theory” section and recounted in detail in Appendices A and B. It is a feature of space-shift extended modeling, and it has no precise analog for ordinary Born modeling.

Ten Kroode (2012) describes the construction of such an operator for 3D extended Kirchhoff modeling (that is, modeling reflections from interfaces, rather than from perturbations in the material parameter fields). Our work is directly inspired by his. The construction described in ten Kroode (2012) approximately inverts a space-shift extended Kirchhoff operator and an angle-domain analog, via the Radon transform relation between space-shift and scattering angle described first by Sava and Fomel (2003). Ten Kroode (2012) suggests that an inverse is possible in the *unfocused* case, in which image energy is spread over subsurface offset (or nontrivially dependent on the scattering angle). Bleistein et al. (2005), Zhang et al. (2007), Zhang and Sun (2008), Xu et al. (2011), and Tang et al. (2013) describe an analogous construction of an inversion operator that produces accurate images and angle-domain extended gathers in the focused case. In all of these works (and in ours), geometric optics/acoustics (ray theory) justifies the inversion construction — but no ray-theory constructions appear in the final result! The critical observation at the root of this remarkable fact is due to Bleistein et al. (2005): the Hessian (“Beylkin”) determinants arising in stationary phase approximation of the normal operator  $\bar{F}^*[v_0]\bar{F}[v_0]$  factor into reciprocal geometric amplitudes and other, ray-independent, terms. These geometric amplitude factors *cancel* the geometric amplitudes present in the propagating fields. The remaining expressions are free of ray-theoretic quantities.

Our results differ in several respects from those cited in the last paragraph. Unlike ten Kroode (2012), we base our construction on Born (rather than Kirchhoff) modeling to produce an approximate extended LSM operator. The underlying mechanism of the inverse construction is somewhat clearer in this case. Also, just as Born modeling is the linearization of the full-waveform modeling operator, extended Born modeling is the linearization of an extended version of full-waveform modeling (Symes, 2008), and the approximate inverse may be useful in accelerating gradient-based, image-domain full-waveform inversion algorithms (Sun and Symes, 2012; Biondi and Almomin, 2014). Although most of the cited work concerns the scattering angle extension, the link between the angle domain and the subsurface offset domain is simply a transform (Sava and Fomel, 2003). There being no intrinsic additional information content in the angle representation, we present our results exclusively in terms of subsurface offset. Unlike much previous work (for instance, Xu et al. [2011]; see the discussion surrounding equation 10), we do not assume that the background velocity  $v_0$  is kinematically correct. Like the Kirchhoff inversion of ten Kroode (2012), our Born inversion operator produces extended models with accurate amplitudes even when the image volume is unfocused. Such accurate extended inversion is critical for the success of image-domain velocity updating schemes (Kern and Symes, 1994; Liu et al., 2013, 2014; Lameloise et al., 2015).

We give implementation details and numerical illustrations for the 2D constant-density acoustics version of an approximate inversion taking the form given in equation 1. Our work seems to be the

first to confirm explicitly, by numerical example, that this operator is actually an inversion of the Born modeling operator  $\bar{F}[v_0]$ , that is, that the output of the inversion operator, input to the modeling operator, reproduces the data with reasonable accuracy.

The form of the approximate inverse (equation 1) with symmetric positive definite weight operators, also seems to be new, and it has a remarkable implication: For norms in model and data spaces given by the weight operators,  $\bar{F}[v_0]$  is the adjoint of  $\bar{F}[v_0]$ . That is, the extended modeling operator is approximately unitary with this choice of norms. The authors have verified the consequent drastic convergence acceleration for extended LSM via Krylov subspace iteration. These results will be reported elsewhere.

As other authors have suggested, a 3D approximate inversion operator may be written in precisely the same form (equation 1), with a virtually identical derivation. However, explanation of the theory is somewhat simpler in 2D, the computation implementation is a good deal simpler, examples are less demanding, and the results are easier to present.

We end this overview with two caveats. First, we have considered only very idealized acquisition geometry (and that in 2D). We have neglected the implications of coarse sampling, more complex source receiver geometry, such as ocean bottom seismometer (OBS) recording, broadband technology, and availability of direct measurements of quantities other than pressure for formulation of RTM-based approximate inversion operators. Others have addressed some of these issues (Tang et al., 2013); some or all will arise in any practical application. Second, although our derivation produces an inverse regardless of focusing, the model-domain weight operator  $W_{\text{model}}[v_0]$  simplifies greatly in the focused case (or equally well for laterally homogeneous velocities), and the examples presented here are limited to these special cases. The “Discussion” section describes the additional steps required to implement full image-volume inversion.

The next section describes space-shift linearized modeling and migration and the construction of the approximate inverse operator. The following section presents several 2D examples and illustrates the features of the approximate inverse mentioned above. We follow the examples with a brief discussion of various possibilities for modification or extension of these results.

## THEORY

In this section, we will first review the concepts of the extended Born modeling operator, its adjoint operator, and their high-frequency approximations. We will then modify the adjoint operator into an approximate inverse operator. Finally, we will discuss implementation details for the inverse operator.

### Extended Born modeling operator and its adjoint

The 2D constant density acoustic-wave equation with causal initial condition is

$$\frac{1}{v^2(\mathbf{x})} \frac{\partial^2 u}{\partial t^2}(\mathbf{x}, t) - \nabla^2 u(\mathbf{x}, t) = f(t, \mathbf{x}, \mathbf{x}_s);$$

$$u(\mathbf{x}, t) \equiv 0, t \ll 0. \quad (2)$$

Here,  $\mathbf{x}$  denotes position within a model of the earth,  $v(\mathbf{x})$  is the acoustic velocity,  $u(\mathbf{x}, t)$  is the acoustic potential, and  $f(t, \mathbf{x}, \mathbf{x}_s)$  is the source term. We assume throughout this paper that  $v$  is

constant in the half-space  $z < 0$ , that is, that  $z = 0$  is an absorbing surface.

In the Born approximation, the velocity model is split into a smooth part  $v_0$  and a singular or oscillatory part  $\delta v$ :

$$v(\mathbf{x}) = v_0(\mathbf{x}) + \delta v(\mathbf{x}). \quad (3)$$

These two parts, respectively, correspond to a smooth long-wavelength, large-scale background model, which will not produce reflections, and a short-wavelength, small-scale perturbation model, which contains all the high-resolution features.

Of course, one can construct a perturbation expansion for any decomposition of  $v$  into two summands; however, the corresponding perturbation approximation to the acoustic field is most accurate when the scales are separated; that is,  $\delta v$  has very small mean over distances, on which  $v_0$  varies significantly (Symes, 2009).

The first-order perturbation in the acoustic-potential field  $\delta u$  corresponding to  $\delta v$  may be expressed in terms of the causal Green's function  $G(\mathbf{x}, \mathbf{y}, t)$  for a given background model  $v_0$ . Restricting  $\delta u$  to the source and receiver positions  $\mathbf{x}_s, \mathbf{x}_r$  results in an integral operator expression for the Born modeling operator  $F[v_0]$ :

$$\begin{aligned} (F[v_0]\delta v)(\mathbf{x}_s, \mathbf{x}_r, t) &= \frac{\partial^2}{\partial t^2} \int d\mathbf{x} d\tau G(\mathbf{x}_s, \mathbf{x}, \tau) \\ &\times \frac{2\delta v(\mathbf{x})}{v_0(\mathbf{x})^3} G(\mathbf{x}, \mathbf{x}_r, t - \tau). \end{aligned} \quad (4)$$

The adjoint operator  $F[v_0]^*$  is the operator implemented by one common variant of RTM. It is applied to a set of data trace perturbations  $\delta d(\mathbf{x}_s, \mathbf{x}_r, t)$  by

$$\begin{aligned} (F^*[v_0]\delta d)(\mathbf{x}) &= \frac{2}{v_0(\mathbf{x})^3} \int d\mathbf{x}_s d\mathbf{x}_r dt d\tau G(\mathbf{x}_s, \mathbf{x}, \tau) \\ &\times G(\mathbf{x}, \mathbf{x}_r, t - \tau) \frac{\partial^2}{\partial t^2} \delta d(\mathbf{x}_s, \mathbf{x}_r, t). \end{aligned} \quad (5)$$

An appropriate version of subsurface offset extended Born modeling introduces dependence of  $\delta v$  (but not  $v_0$ ) on an additional parameter,  $\mathbf{h}$ , essentially the offset between the sunken source and sunken receiver in Claerbout's (1985) survey-sinking imaging condition (Symes, 2008; Stolk et al., 2009b). In terms of Green's functions, the subsurface extended Born modeling operator and its adjoint (applied to a data perturbation  $\delta d$ ) are

$$\begin{aligned} (\bar{F}[v_0]\delta \bar{v})(\mathbf{x}_s, \mathbf{x}_r, t) &= \frac{\partial^2}{\partial t^2} \int d\mathbf{x} d\mathbf{h} d\tau G(\mathbf{x}_s, \mathbf{x} - \mathbf{h}, \tau) \\ &\times \frac{2\delta \bar{v}(\mathbf{x}, \mathbf{h})}{v_0(\mathbf{x})^3} G(\mathbf{x} + \mathbf{h}, \mathbf{x}_r, t - \tau), \end{aligned} \quad (6)$$

$$\begin{aligned} (\bar{F}^*[v_0]\delta d)(\mathbf{x}, \mathbf{h}) &= \frac{2}{v_0(\mathbf{x})^3} \int d\mathbf{x}_s d\mathbf{x}_r dt d\tau G(\mathbf{x}_s, \mathbf{x} - \mathbf{h}, \tau) \\ &\times G(\mathbf{x} + \mathbf{h}, \mathbf{x}_r, t - \tau) \frac{\partial^2}{\partial t^2} \delta d(\mathbf{x}_s, \mathbf{x}_r, t). \end{aligned} \quad (7)$$

In Claerbout's (1985) original conception, the subsurface offset  $\mathbf{h}$  is horizontal. Ten Kroode (2012) also adopts this convention, and we

follow it here. Thus, we write  $h$  rather than  $\mathbf{h}$  for the (scalar) horizontal subsurface offset in 2D.

### High-frequency approximation

The progressing wave approximation (Courant and Hilbert, 1962) of the Green's function is

$$G(\mathbf{x}_s, \mathbf{x}, t) \simeq a(\mathbf{x}_s, \mathbf{x}) S(t - T(\mathbf{x}_s, \mathbf{x})). \quad (8)$$

In equation 8, the amplitude  $a(\mathbf{x}_s, \mathbf{x})$  and the traveltime  $T(\mathbf{x}_s, \mathbf{x})$  solve the transport and eikonal equation, respectively, and  $S(t)$  is a singular, causal waveform, the choice of which depends on the space dimension. Approximation 8 is only valid locally, between the source point and the nearest caustic or conjugate (multipath) point. The conclusions we draw below are valid more globally; however, provided that the traveltime injectivity condition holds: A two-way traveltime along a reflected ray pair determines the one-way traveltimes of source and receiver rays. Ten Kroode (2012) gives a detailed justification for the global validity of similar conclusions in the 3D case. We confine ourselves in this paper to numerical evidence for global 2D results.

In the 2D case, the leading singularity is proportional to the generalized function  $S(t) = t_+^{-1/2} = t^{-1/2} H(t)$ . Replacing the Green's function by the progressing wave approximation 8 in the expression 6 for the extended Born modeling operator and using the identity (Gel'fand and Shilov, 1958)

$$t_+^{-1/2} * t_+^{-1/2} = \left( \Gamma\left(\frac{1}{2}\right) \right)^2 H(t) = \pi H(t), \quad (9)$$

we obtain

$$(\bar{F}[v_0]\delta \bar{v})(\mathbf{x}_s, \mathbf{x}_r, t) \simeq \frac{\partial}{\partial t} \int d\mathbf{x} d\mathbf{h} a_s a_r \delta(t - T_s - T_r) \frac{2\pi \delta \bar{v}(\mathbf{x}, h)}{v_0(\mathbf{x})^3}, \quad (10)$$

in which we have denoted amplitudes  $a(\mathbf{x}_s, \mathbf{x} - h)$ ,  $a(\mathbf{x} + h, \mathbf{x}_r)$  as  $a_s, a_r$  and traveltime  $T(\mathbf{x}_s, \mathbf{x} - h)$ ,  $T(\mathbf{x} + h, \mathbf{x}_r)$  as  $T_s, T_r$ . We can also give the same treatment to the migration operator:

$$(\bar{F}[v_0]^* \delta d)(\mathbf{x}, h) \simeq \frac{2\pi}{v_0(\mathbf{x})^3} \int d\mathbf{x}_s d\mathbf{x}_r a_s a_r \frac{\partial}{\partial t} \delta d(\mathbf{x}_s, \mathbf{x}_r, T_s + T_r). \quad (11)$$

### Asymptotic inverse operator

The derivation of the inverse operator starts from understanding the high-frequency leading order behavior of the normal operator  $\bar{F}^* \bar{F}$ . A bit of foresight and simplification of the computations suggest examining the modified normal operator  $(I_t \bar{F})^* (I_t \bar{F})$  instead, with  $I_t$  the causal indefinite time integration operator, inverse to  $\partial/\partial t$ . Combining equations 10 and 11 yields an integral representation for  $(I_t \bar{F})^* (I_t \bar{F})$ . Appendix A sketches a lengthy, but standard, computation based on the principle of stationary phase, showing that this integral representation has an asymptotic (high-frequency and short-scale) approximation in the form of an oscillatory integral:

$$(I_t \bar{F})^* (I_t \bar{F}) \delta \bar{v}(x, z, h) \approx - \int dk_x dk_z dk_h e^{i(k_x x + k_h h + k_z z)} \hat{\delta v}(k_x, k_h, k_z) \times \left[ \frac{2\pi v_0^{-5}}{k_{xz} k_{hz}} P a_r^2 a_s^2 \left( \frac{\partial \alpha_s}{\partial x_s} \frac{\partial \alpha_r}{\partial x_r} \right)^{-1} \right]. \quad (12)$$

The integrand factor in brackets will be explained below: If it were identically = 1, then up to nuisance factors the right side would recover the velocity perturbation  $\delta \bar{v}$ , accurately at short-length scales.

The remainder of this section explains how the factor in brackets simplifies, all ray-trace-dependent quantities cancel, additional filtering and scaling operators applied to data (input) and model (output) lead to an approximate identity, and precisely how it comes to have the form of the main result (equation 1).

The integrand in equation 12 is a function of the extended phase space variables  $(x, z, h, k_x, k_z \text{ and } k_h)$ . Several of its factors are related to source and receiver rays traced from  $(x \pm h, z)$  to the surface. These include the following:

- 1) Source and receiver geometric amplitudes (spreading factors)  $a_s$  and  $a_r$ ;
- 2) rates of change of receiver ray angle (with the vertical)  $\alpha_r$  with respect to receiver coordinate  $x_r$ ;
- 3) rates of change of source ray angle (with the vertical)  $\alpha_s$  with respect to source coordinate  $x_s$ .

The initial slowness vectors  $\mathbf{k}_s$  of the source ray (at  $(x - h, z)$ ) and  $\mathbf{k}_r$  of the receiver ray (at  $(x + h, z)$ ) solve the system of equations

$$\begin{aligned} k_x &= \frac{k_x^r + k_x^s}{2}, \\ k_h &= \frac{k_x^r - k_x^s}{2}, \\ k_z &= \frac{k_z^r + k_z^s}{2}, \end{aligned}$$

$$v(x + h, z)^2 ((k_x^r)^2 + (k_z^r)^2) = v(x - h, z)^2 ((k_x^s)^2 + (k_z^s)^2). \quad (13)$$

The first three conditions in the system (equation 13) follow from the stationary phase conditions (equation A-6), as explained in Appendix A. The last condition expresses the equality of temporal frequency along the source and receiver rays. These conditions together amount to a version of Snell's law appropriate for space-shift extended modeling.

Rays with initial conditions  $(x + h, z, k_x^r, k_z^r)$  and  $(x - h, z, k_x^s, k_z^s)$  intersect source and receiver datums at points  $(x_s, z_s)$  and  $(x_r, z_r)$ , thus making  $x_r, x_s$  functions of the extended phase space variables  $(x, z, h, k_x, k_z, k_h)$ , and therefore also the geometric amplitudes  $a_r$  and  $a_s$ . The ray angles are as well because, for instance,  $\tan \alpha_r = k_x^r / k_z^r$ ; therefore, so are their derivatives with respect to source and receiver coordinates.

The other factors in the bracketed integrand factor in equation 12 are *explicit, algebraic* functions of the phase variables, whose definition does not require ray tracing at all. The  $(x, z)$  and  $(h, z)$  wave-numbers are defined as

$$k_{xz} = (k_x^2 + k_z^2)^{\frac{1}{2}}, \quad k_{hz} = (k_h^2 + k_z^2)^{\frac{1}{2}}. \quad (14)$$

The remaining factor  $P$  is homogeneous of degree zero in  $(k_x, k_z, k_h)$ , and it depends algebraically on these frequency variables

and on  $v(x - h, z)$ ,  $v(x, z)$  and  $v(x + h, z)$ . A full definition of  $P$  is given in Appendix A. For now, note that  $P = 1$ , when  $h = 0$ .

Next, we invoke the result of Appendix B, relating geometric amplitudes to angular rates of change:

$$\begin{aligned} a_r^2 &= \frac{1}{8\pi^2} \frac{v_r}{\cos \theta_r} \frac{d\alpha_r}{dx_r}, \\ a_s^2 &= \frac{1}{8\pi^2} \frac{v_s}{\cos \theta_s} \frac{d\alpha_s}{dx_s}. \end{aligned} \quad (15)$$

In these expressions,  $\theta_r$  and  $\theta_s$  are arrival angles of receiver and source rays at receiver and source. Because the receiver and source ray data are functions of the phase variables, so are  $\theta_r$  and  $\theta_s$ .

Remarkably, the amplitudes and angular rates of change in equation 12 *cancel* due to equation 15. This cancellation of both geometric amplitudes occurs only for extended modeling. The analogous computation for single shot records, for example, leads to cancellation of the receiver ray amplitude  $a_r$  only. In that setting, inversion requires an additional operator, essentially Claerbout's division imaging condition (Stolk et al., 2009a), to compensate for the source amplitude field. For the horizontal offset extension, the additional integrations over the subsurface offset variable lead to an additional  $\alpha_s$  derivative via the stationary phase, thus canceling the source amplitude as well. See Xu et al. (2011) for a version of this argument using the scattering angle rather than the subsurface offset.

Taking advantage of this observation and simplifying, equation 12 becomes

$$(I_t \bar{F})^* (I_t \bar{F}) \delta \bar{v}(x, z, h) \approx - \int dk_x dk_z dk_h e^{i(k_x x + k_h h + k_z z)} \hat{\delta v}(k_x, k_h, k_z) \times \left[ \frac{v_0^{-5}}{32\pi^3 k_{xz} k_{hz}} P \frac{v_s}{\cos \theta_s} \frac{v_r}{\cos \theta_r} \right]. \quad (16)$$

The integrand on the right side still appears entangled with ray-theoretic constructions, namely, the arrival angles  $\theta_r, \theta_s$ . To eliminate these, a further modification of the normal operator is necessary. The expression 10 for the modeling operator using asymptotic Green's functions implies that

$$\begin{aligned} \frac{\partial}{\partial z_r} I_t \bar{F}[v_0] \delta \bar{v}(\mathbf{x}_s, \mathbf{x}_r, t) &\approx \\ &- \int d\mathbf{x} d\mathbf{h} a_s a_r \frac{\partial T_r}{\partial z_r} \frac{\partial \delta}{\partial t} (t - T_s - T_r) \frac{2\pi \delta \bar{v}(\mathbf{x}, h)}{v_0(\mathbf{x})^3} \\ &= - \frac{\partial}{\partial t} \int d\mathbf{x} d\mathbf{h} a_s a_r \frac{\partial T_r}{\partial z_r} \delta(t - T_s - T_r) \frac{2\pi \delta \bar{v}(\mathbf{x}, h)}{v_0(\mathbf{x})^3}. \end{aligned} \quad (17)$$

From the eikonal equation,

$$\frac{\partial T_r}{\partial z_r} = - \frac{\cos \theta_r}{v_r}. \quad (18)$$

Combine equations 17 and 18, use the abbreviations  $D_{z_r} = \partial / \partial z_r$ , and cancel the two time derivatives with time integrations to obtain

$$I_t D_{z_r} I_t \bar{F}[v_0] \delta \bar{v}(\mathbf{x}_s, \mathbf{x}_r, t) \approx \int d\mathbf{x} d\mathbf{h} a_s a_r \frac{\cos \theta_r}{v_r} \delta(t - T_s - T_r) \frac{2\pi \delta \bar{v}(\mathbf{x}, h)}{v_0(\mathbf{x})^3}. \quad (19)$$

Precisely, the same manipulations with the source-related quantities lead to

$$I_t D_{z_s} I_t D_{z_r} I_t \bar{F}[v_0] \delta \bar{v}(\mathbf{x}_s, \mathbf{x}_r, t) \approx \int d\mathbf{x} d\mathbf{h} a_s a_r \frac{\cos \theta_r \cos \theta_s}{v_r v_s} \delta(t - T_s - T_r) \frac{2\pi \delta \bar{v}(\mathbf{x}, h)}{v_0(\mathbf{x})^3}. \quad (20)$$

The right side of equation 20 defines an operator differing from  $I_t \bar{F}[v_0]$  only in having additional receiver- and source-dependent factors multiplying the amplitudes. Therefore, the stationary phase computations in Appendix A combined with the amplitude-angle relations of Appendix B lead to an asymptotic approximation similar to equation 16:

$$(I_t \bar{F})^* (I_t D_{z_s} I_t D_{z_r} I_t \bar{F}) \delta \bar{v}(x, z, h) \approx - \int dk_x dk_z dk_h e^{i(k_x x + k_h h + k_z z)} \hat{\delta v}(k_x, k_h, k_z) \times \left[ \frac{v_0^{-5}}{32\pi^3 k_{xz} k_{hz}} P \right]. \quad (21)$$

The source wavefield is downgoing, and the receiver wavefield is upcoming (again, we assume a homogeneous velocity in  $z < \max(z_s, z_r)$ ). Therefore, if  $d = \bar{F}[v_0] \delta \bar{v}$  is in the range of the extended Born modeling operator (this observation is due to [ten Kroode, 2012](#)), then

$$I_t D_{z_r} I_t d(x_r, t; x_s) = \frac{1}{8\pi^3} \int dk_s dk_r d\omega \exp(i(k_s x_s + k_r x_r + \omega t)) \times \left( + \sqrt{\frac{1}{v_0^2} - \left(\frac{k_r}{\omega}\right)^2} \right) \hat{I}_t d(k_r, \omega; k_s). \quad (22)$$

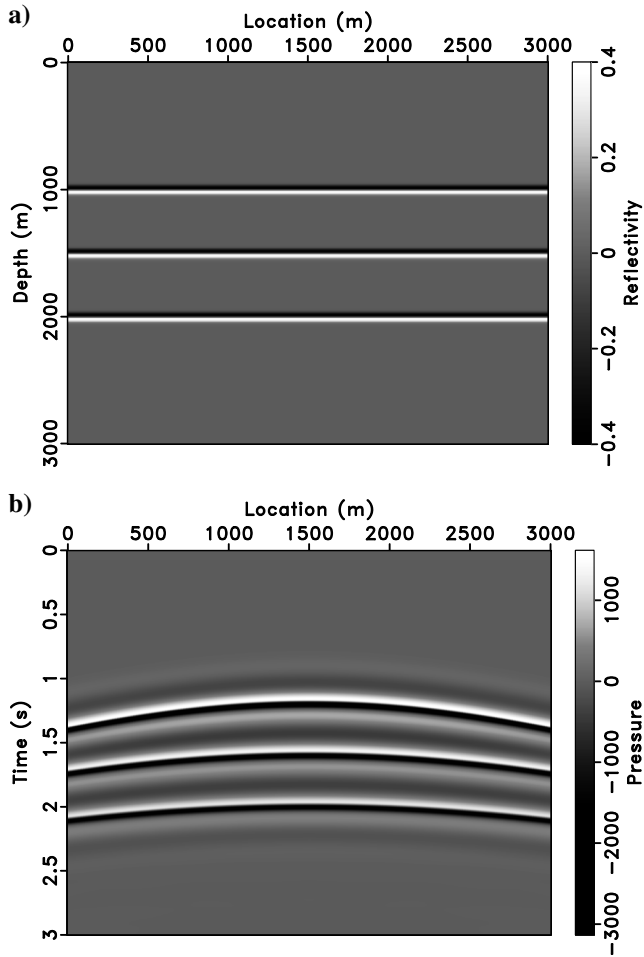


Figure 1. (a) Reflectivity model ( $\delta v$ ) with a constant background model ( $v_0 = 2500$  m/s). (b) One-shot ( $x_s = 1500$  m) simulated Born data.

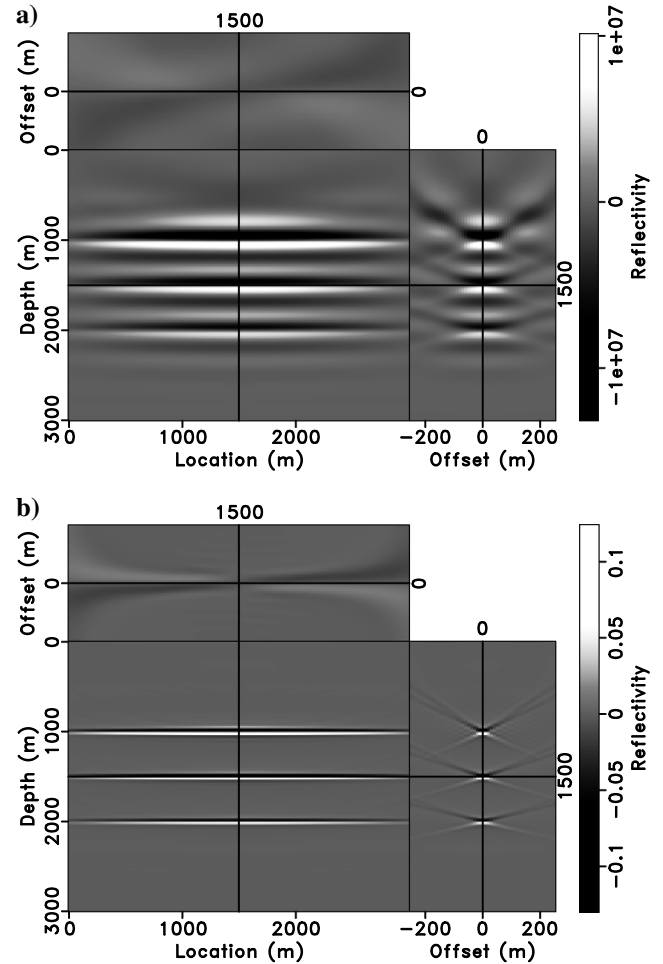


Figure 2. (a) Extended RTM image. (b) Extended inverted reflectivity model.



$$\begin{aligned}
& I_t D_{z_s} I_t d(x_r, t; x_s) \\
&= \frac{1}{8\pi^3} \int dk_s dk_r d\omega \exp(i(k_s x_s + k_r x_r + \omega t)) \\
&\times \left( -\sqrt{\frac{1}{v_0^2} - \left(\frac{k_s}{\omega}\right)^2} \right) \widehat{I_t d}(k_r, \omega; k_s). \quad (23)
\end{aligned}$$

It follows from equations 22 and 23 that the operator  $-(I_t)^*(I_t D_{z_s})(I_t D_{z_r})I_t$  has the same effect on data output by the forward map  $\bar{F}[v_0]$  as the positive definite symmetric operator  $W_{\text{data}}[v_0]$ , defined by

$$\begin{aligned}
W_{\text{data}}[v_0] d(x_r, t; x_s) &= (I_t)^* \frac{1}{8\pi^3} \int dk_s dk_r d\omega \\
&\times \exp(i(k_s x_s + k_r x_r + \omega t)) \left( \sqrt{\frac{1}{v_0^2} - \left(\frac{k_r}{\omega}\right)^2} \right) \\
&\times \left( \sqrt{\frac{1}{v_0^2} - \left(\frac{k_s}{\omega}\right)^2} \right) (\widehat{I_t d})(k_r, \omega; k_s). \quad (24)
\end{aligned}$$

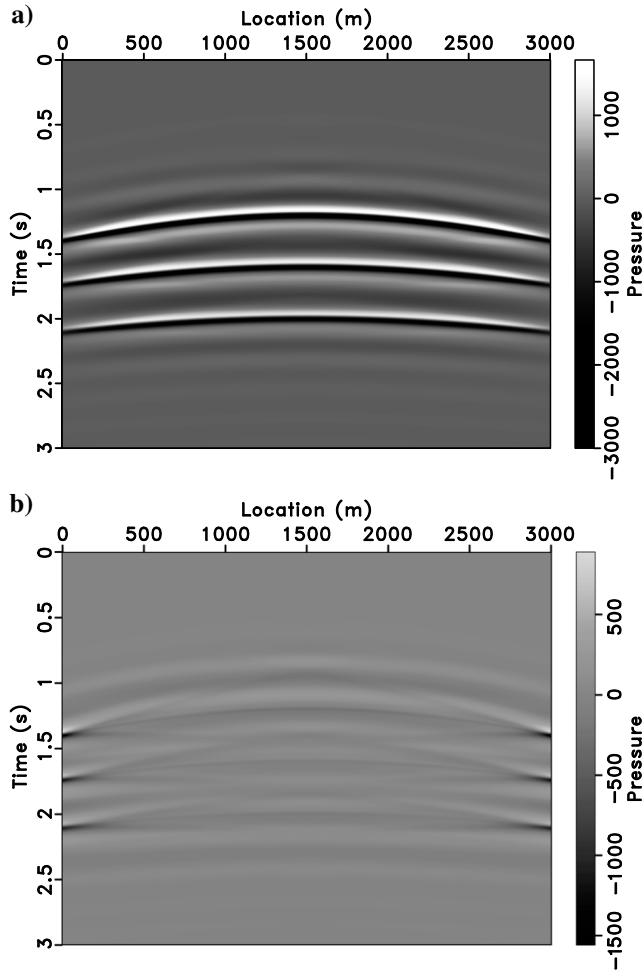


Figure 3. (a) Resimulated data of the inverted reflectivity model. (b) Difference between the resimulated data and original data.

Explicitly,

$$W_{\text{data}}[v_0] \bar{F}[v_0] = -(I_t)^* I_t D_{z_s} I_t D_{z_r} I_t \bar{F}[v_0]. \quad (25)$$

Note that  $W_{\text{data}}[v_0]$  depends only on the values of  $v_0$  near the sources and receivers: It is completely independent of the behavior of  $v_0$  for  $z > \max(z_r, z_s)$ .

Oscillatory integral operators of the type appearing on the right side of equation 21 have come to be called *pseudodifferential*, and have a number of important properties, of which we must use several. For example, the product of two such operators is another such: The Fourier representation amplitude (or *symbol*) of the product is the product of the symbols of the operator factors, up to an asymptotically negligible error. It follows that such operators (on scalar functions) commute up to an asymptotically negligible error. An operator with a nonvanishing symbol is asymptotically invertible, and the symbol of the inverse is the reciprocal of the symbol. See, for instance, Taylor (1981) for an account of the calculus of pseudodifferential operators.

It follows that we can write the right side of equation 21 as

$$\approx (W_{\text{model}}[v_0]^{-1} \delta \bar{v})(x, z, h), \quad (26)$$

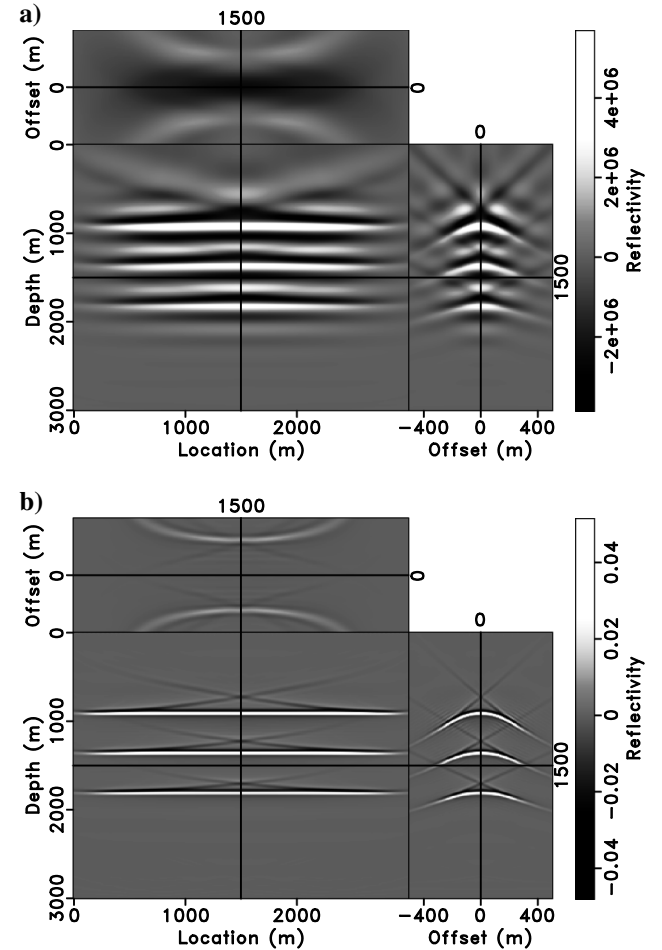


Figure 4. (a) Extended RTM image. (b) Extended inverted reflectivity model, both using an incorrect background velocity model

in which

$$W_{\text{model}}[v_0]^{-1}u(x, z, h) = \frac{1}{8\pi^3} \int dk_x dk_h dk_z \times \exp(i(k_x x + k_h h + k_z z)) \frac{v_0(x, z)^{-5}}{4k_{xz}k_{hz}} \times P(x, z, h, k_x, k_z, k_h) \hat{u}(k_x, k_z, k_h). \quad (27)$$

As follows from the facts mentioned above

$$W_{\text{model}}[v_0]u(x, z, h) = \frac{1}{8\pi^3} \int dk_x dk_h dk_z \times \exp(i(k_x x + k_h h + k_z z)) \frac{4v_0(x, z)^5 k_{xz} k_{hz}}{P(x, z, h, k_x, k_z, k_h)} \times \hat{u}(k_x, k_z, k_h). \quad (28)$$

Combining equations 21, 25, and 26 establishes the main conclusion of our paper, equation 1, with  $W_{\text{model}}$  defined in equation 28 and  $W_{\text{data}}$  defined in equation 24.

We end this section by describing how the approximate inverse operator 1 defines an approximate (nonextended) LSM. Given a velocity perturbation  $\delta v(x, z)$ , the corresponding extended model is

$$\delta \bar{v}(x, z, h) = \delta v(x, z) \delta(h). \quad (29)$$

Let

$$\delta d = F[v_0] \delta v \quad (30)$$

be the corresponding Born data. Then,

$$\delta \bar{v} = W_{\text{model}}[v_0] \bar{F}[v_0]^* W_{\text{data}}[v_0] \delta d, \quad (31)$$

whence,

$$\delta v(x, z) = \int dh \phi(h) (W_{\text{model}}[v_0] \bar{F}[v_0]^* W_{\text{data}}[v_0] \delta d)(x, z, h), \quad (32)$$

for any weight function  $\phi(h)$  satisfying  $\phi(0) = 1$ .

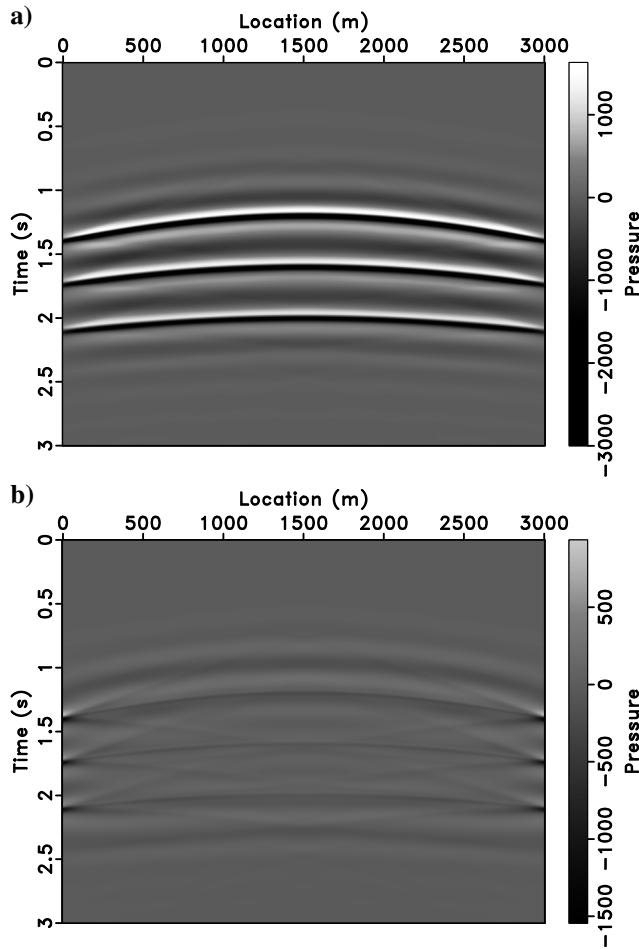


Figure 5. (a) Resimulated data of the inverted reflectivity model using an incorrect background velocity. (b) Difference between the resimulated data and original data.

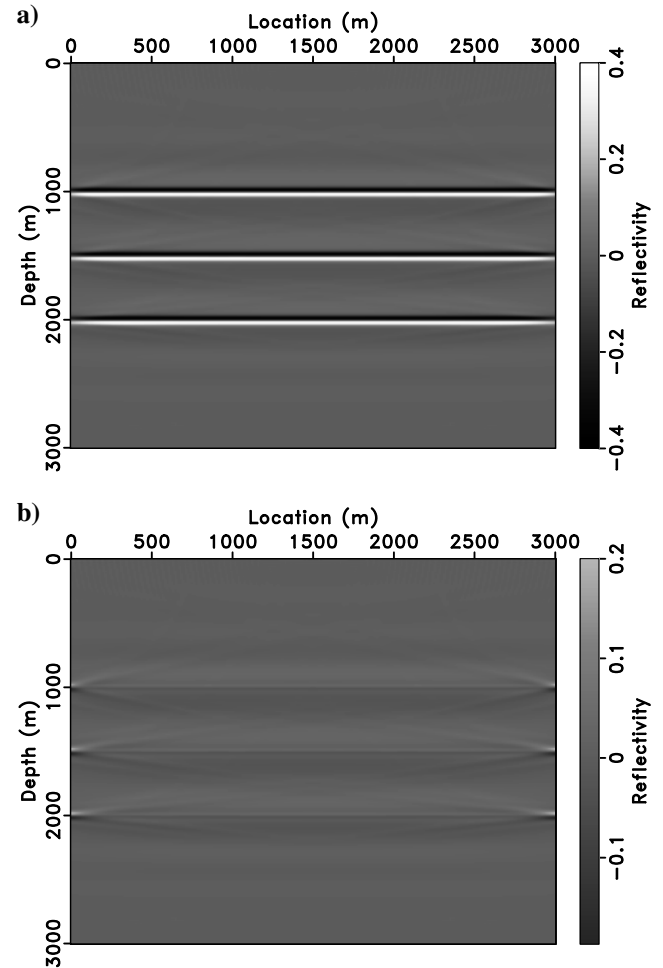


Figure 6. (a) Nonextended inverted reflectivity model ( $\sum_h i(x, h)$ , where  $i(x, h)$  is the extended inversion result). (b) Difference between the nonextended inversion result and the original reflectivity model.

The arbitrariness of the weight function  $\phi(h)$ , subject only to the constraint  $\phi(0) = 1$ , might seem strange. Viewing the formula 29 in terms of the related angle-domain image volume, as in Sava and

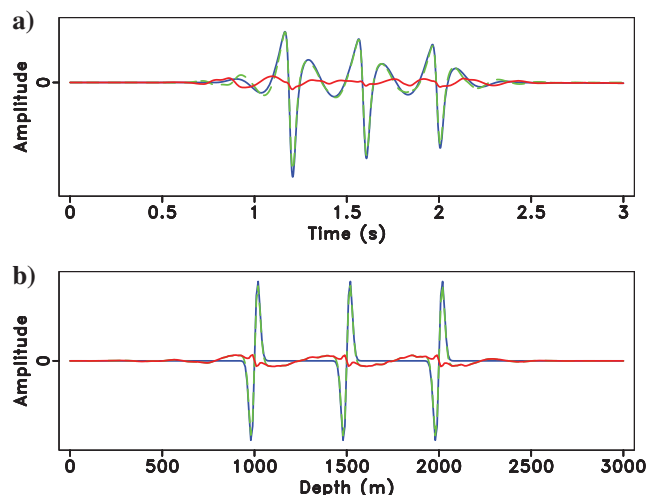


Figure 7. (a) One-trace comparison ( $x = 1500$  m) between the observed data (solid blue line) and predicted data from the inverted reflectivity model (dashed green line). The difference is shown as the solid red line. (b) One-trace comparison ( $x = 1500$  m) between the reflectivity model (solid blue line) and the nonextended inverted reflectivity model (dashed green line). The difference is shown as the solid red line.

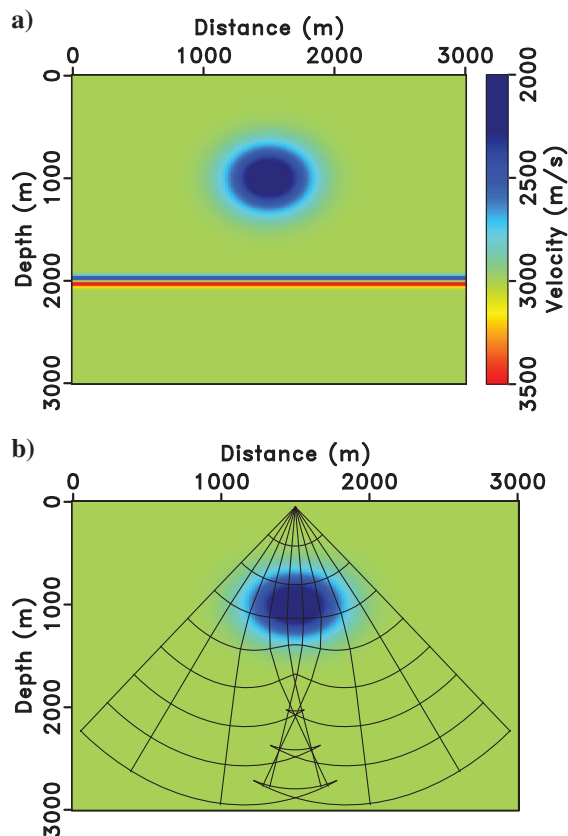


Figure 8. (a) Gaussian lens background velocity model with a reflector at 2 km. (b) The rays and wavefronts in the Gaussian lens velocity model.

Fomel (2003), suggests an alternate meaning for this formula. Via the Radon transform, a weighted average over the offset is equivalent to a weighted average over angles of the corresponding angle-domain volume. A weight function  $\phi(h)$  spread uniformly over the offset range, as we have used in the examples of the next section, corresponds to an angle-domain weight function concentrated near zero (scattering angle). For well-focused noise-free data, inversion using a small range of scattering angles should be reasonably accurate, as indeed the examples presented in the next section attest. On the other hand, a choice of  $\phi(h)$  concentrated near  $h = 0$  would correspond to estimating  $\delta v$  as a stack over a wide range of scattering angles. One might expect the estimate so obtained to be less sensitive to incoherent or numerical noise.

### Implementation details

In the implementation of the asymptotic inverse operator, any suitable time- or frequency-domain method can be used to approximate  $\bar{F}$  and  $\bar{F}^*$ . We have used a time-domain centered difference scheme of order 2 in time and 8 in space to solve the acoustic wave equation and the well-known adjoint state method (Plessix, 2006) to approximate  $\bar{F}^*$ .

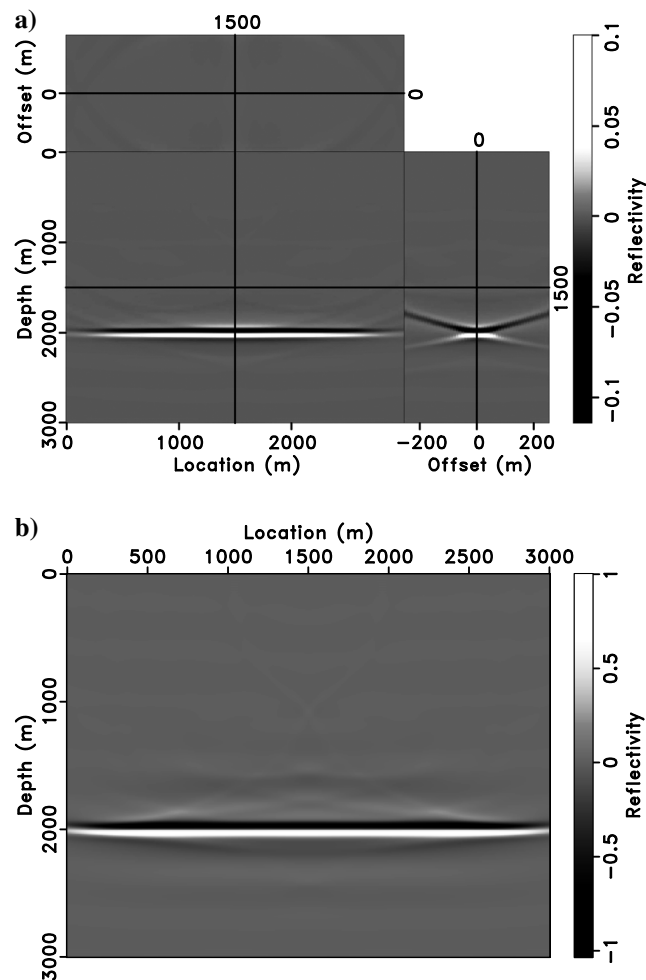


Figure 9. (a) Extended inverted reflectivity model. (b) Nonextended inverted reflectivity model.



For  $W_{\text{data}}$ , equation 25 shows that for model-consistent data, either the definition 24 in terms of one-way operators, or the equivalent expression in terms of source and receiver vertical derivatives (dipole source/receiver) produces the same result, at least in principle. In the numerical experiments reported in the next section, we have chosen the dipole approach. To avoid explicitly computing dipole responses, we have used a trick available for streamer geometry with free surface and relatively shallow tow depth: The ghost sources and receivers automatically supply scaled dipoles. Assuming the source tow depth to be  $z_s$ , the free-surface Green's function  $G_{\text{free}}$  is related to the full-space (absorbing boundary) Green's function  $G$  by

$$G_{\text{free}}(x, z, t; x_s, z_s) = G(x, z, t; x_s, z_s) - G(x, z, t; x_s, -z_s) \approx 2z_s D_{z_s} G(x, z, t; x_s, 0). \quad (33)$$

By reciprocity, a similar approximation applies to the receiver. In application, if  $\bar{F}[v_0]$  is computed with absorbing boundary, then  $\bar{F}[v_0]^*$  can be calculated with free surface, or vice versa — in either case, with appropriate inclusion of  $I_t$  factors, an approximation to  $W_{\text{data}}$  for the absorbing surface problem ensues.

This approximation proved quite convenient and produced consistent results as the central finite-difference implementation (Hou

and Symes, 2014) for a first round of numerical experiments. It might even be applied to actual streamer data with shallow and uniform tow depth. However, it is only good to perhaps half of the notch frequency, so it strictly limits resolution. For more or less arbitrary but uniform source and receiver depths and sampling, the one-way propagator construction (equation 24) would be preferable. As mentioned in the introduction, more sophisticated streamer geometry, OBS recording, and less favorable sampling all would require modification of even the 3D version of our computations.

Application of the weight operator  $W_{\text{model}}[v_0]$  in principle requires the evaluation of an oscillatory integral (equation 28). However, in two special cases, the factor  $P$  may be ignored: As established in Appendix A,  $P = 1$ , if either  $h = 0$  or if  $v_0$  is independent of  $x$ . The first case applies to approximate LSM: If the model and data are consistent, then the image energy focuses at  $h = 0$  and the values of  $P$  for nonzero  $h$  do not contribute to the leading order in frequency (more technically, this is the *pseudolocal* property of  $\bar{F}[v_0]^* W_{\text{data}}[v_0] \bar{F}[v_0]$  and  $W_{\text{model}}[v_0]$ : Both are pseudo-differential, and the image of a physically consistent input is asymptotically negligible away from  $h = 0$ ).

Accordingly, we have set  $P = 1$  in our examples, which fall into one of these two classes. Then,

$$W_{\text{model}}[v_0] \sim 4v_0^5 L, \quad (34)$$

where  $L$  is the filter defined in the Fourier domain by  $k_{xz}k_{hz}$ . In other words,

$$L = \sqrt{-\nabla_{x,z}^2} \sqrt{-\nabla_{h,z}^2}. \quad (35)$$

With these approximations to  $W_{\text{data}}$  and  $W_{\text{model}}$ , the computational cost of the approximate inverse operator  $\bar{F}[v_0]^\dagger$  very similar to that of the extended RTM operator  $\bar{F}[v_0]^*$ .

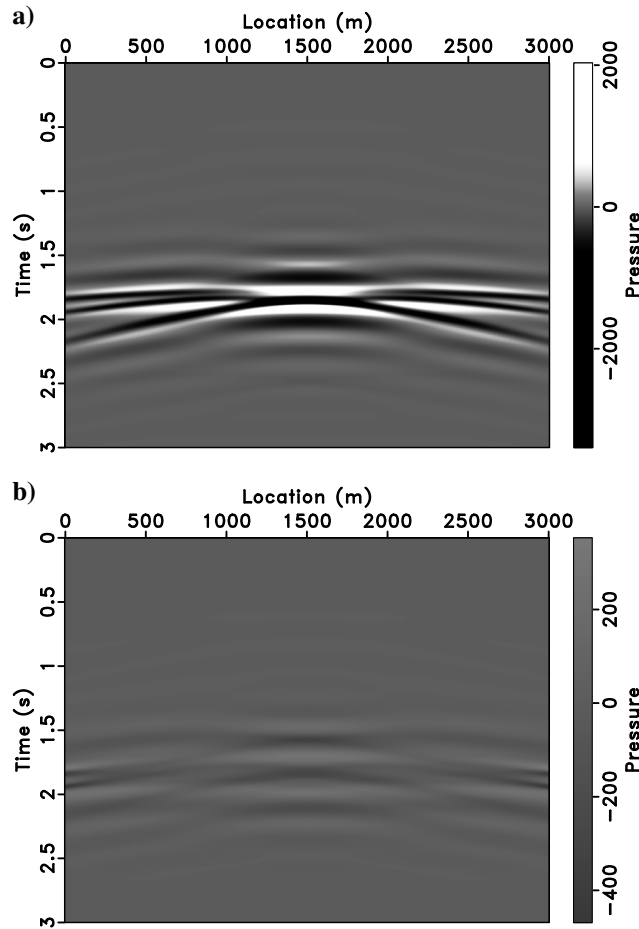


Figure 10. (a) One-shot ( $x_s = 1500$  m) resimulated Born data. (b) Difference between the resimulated data and original data.

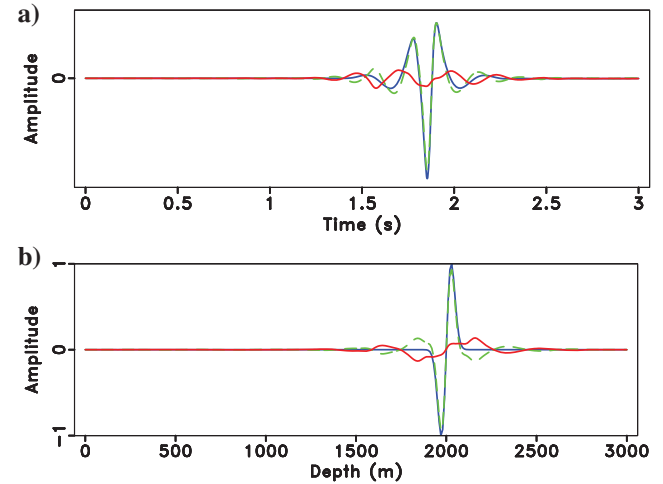


Figure 11. (a) One-trace comparison ( $x = 1500$  m) between the observed data (solid blue line) and predicted data from the inverted reflectivity model (dashed green line). The difference is shown as the solid red line. (b) One-trace comparison ( $x = 1500$  m) between the reflectivity model (solid blue line) and nonextended inverted reflectivity model (dashed green line). The difference is shown as the solid red line.

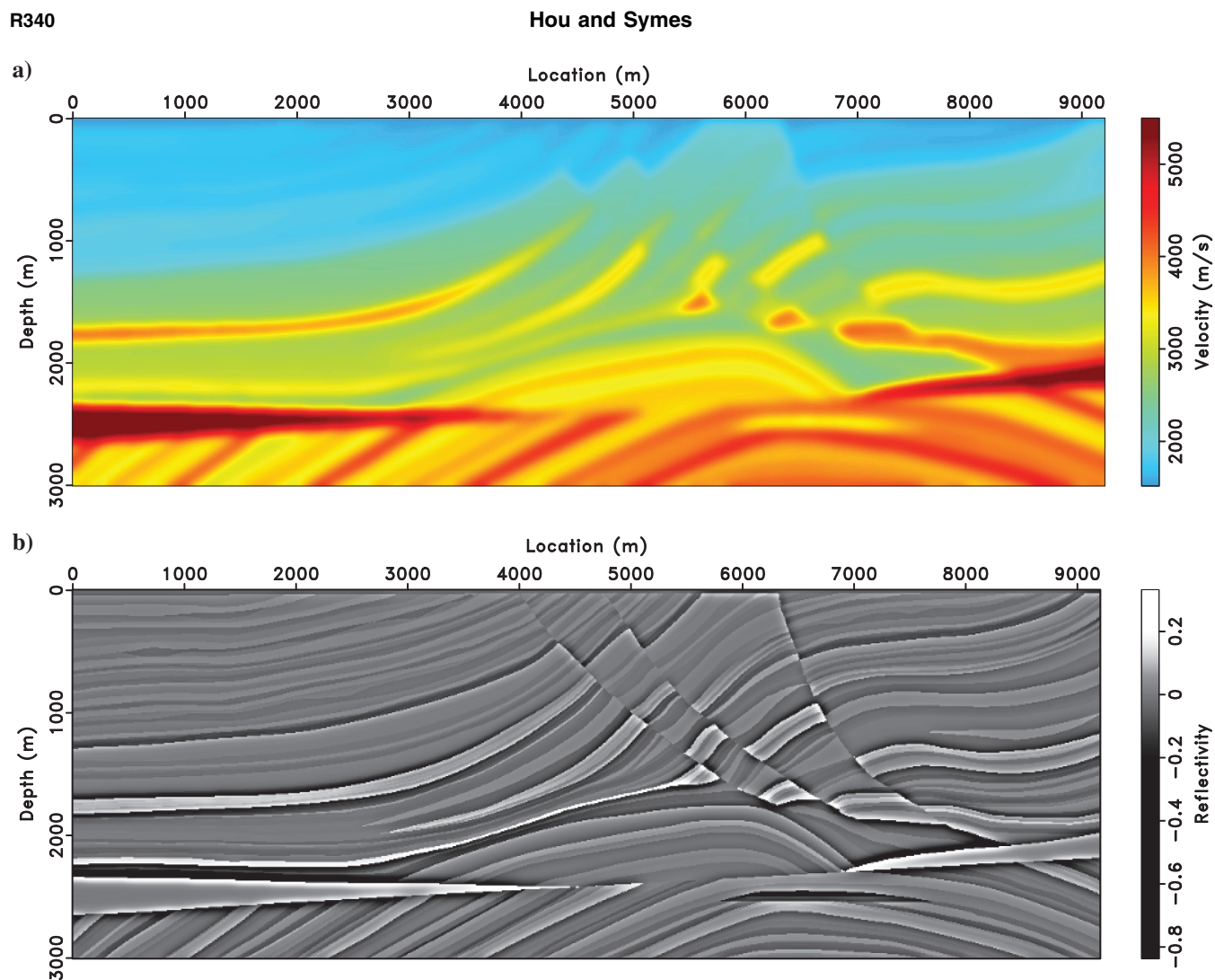


Figure 12. (a) Smoothed background velocity model. (b) Reflectivity model.

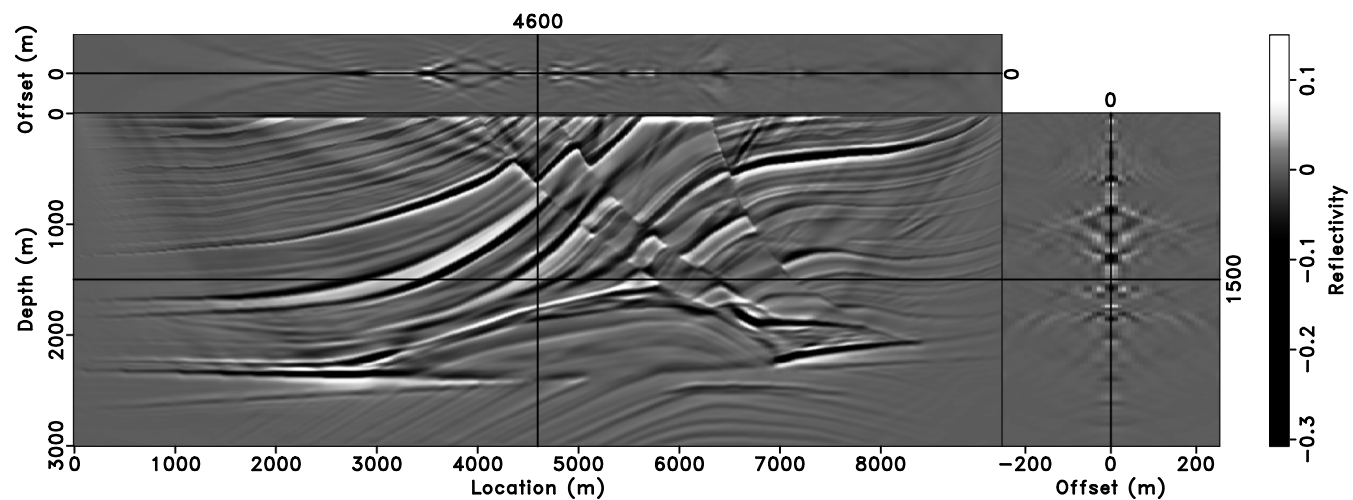


Figure 13. Extended inverted reflectivity model.

## NUMERICAL EXAMPLES

In this section, we will use three numerical examples to illustrate the effectiveness of the inverse operator.

The first model, shown in Figure 1a, combines three flat reflectors at  $z = 1, 1.5, 2$  km with a constant (2500 m/s) background velocity model. The spatial sampling interval of the model is 10 m for  $x$ - and  $z$ -axes. A (2.5–5–35–40)-Hz band-pass wavelet with a 1-ms time interval is used to simulate the Born data (2–8 finite-difference scheme). The 76 shots are evenly spread on the surface ( $z = 0$ ) every 40 m. All of the shots will be recorded by 301 receivers deployed every 10 m on the surface.

The one-shot Born data ( $x_s = 1500$  m) shown in Figure 1b is calculated using equation 6. The extended RTM (equation 7) and the new inverse operator (equation 1) are applied on the Born data. Comparing the migrated image (Figure 2a) and the inverted reflectivity model (Figure 2b), we can clearly see that the inverse operator can focus the energy much better than does extended RTM. It has many LSM qualities, such as improved amplitudes and a tighter wavelet. Thus, the inverse operator at least plays the role of space deconvolution. However, we can never recover the reflector per-

fectly due to the lack of low-frequency data. The inverted reflectivity model will not be a good standard for the effectiveness of the inverse operator. A good way to evaluate the inverse operator would be to compare the “observed” data of the true model (Figure 1b) and the resimulated data of inverted reflectivity model (Figure 3a). The comparisons (Figure 3) show that the data resimulated from the inverted reflectivity model are almost the same as the original data. The degree of approximation shows clearly in the comparison of the middle traces of the original and resimulated data (Figure 7a).

Another point worth noting is that the inverse operator is valid; that is, it produces a data-fitting model whether the background velocity is correct or not. The same comparison between the extended RTM and the inverse operator has been carried out with the incorrect background velocity model (90% of true velocity; see Figure 4). The comparison indicates the above analysis is true even in the presence of velocity error. The resimulation is displayed in Figure 5a, which should be compared with Figure 1b. The difference appears in Figure 5b. Some divergence is unavoidable near the boundary as a result of the acquisition geometry. Apart from that, the resimulation matches the original Born data extremely well. We conclude that the operator defined in equation 1 is an accurate

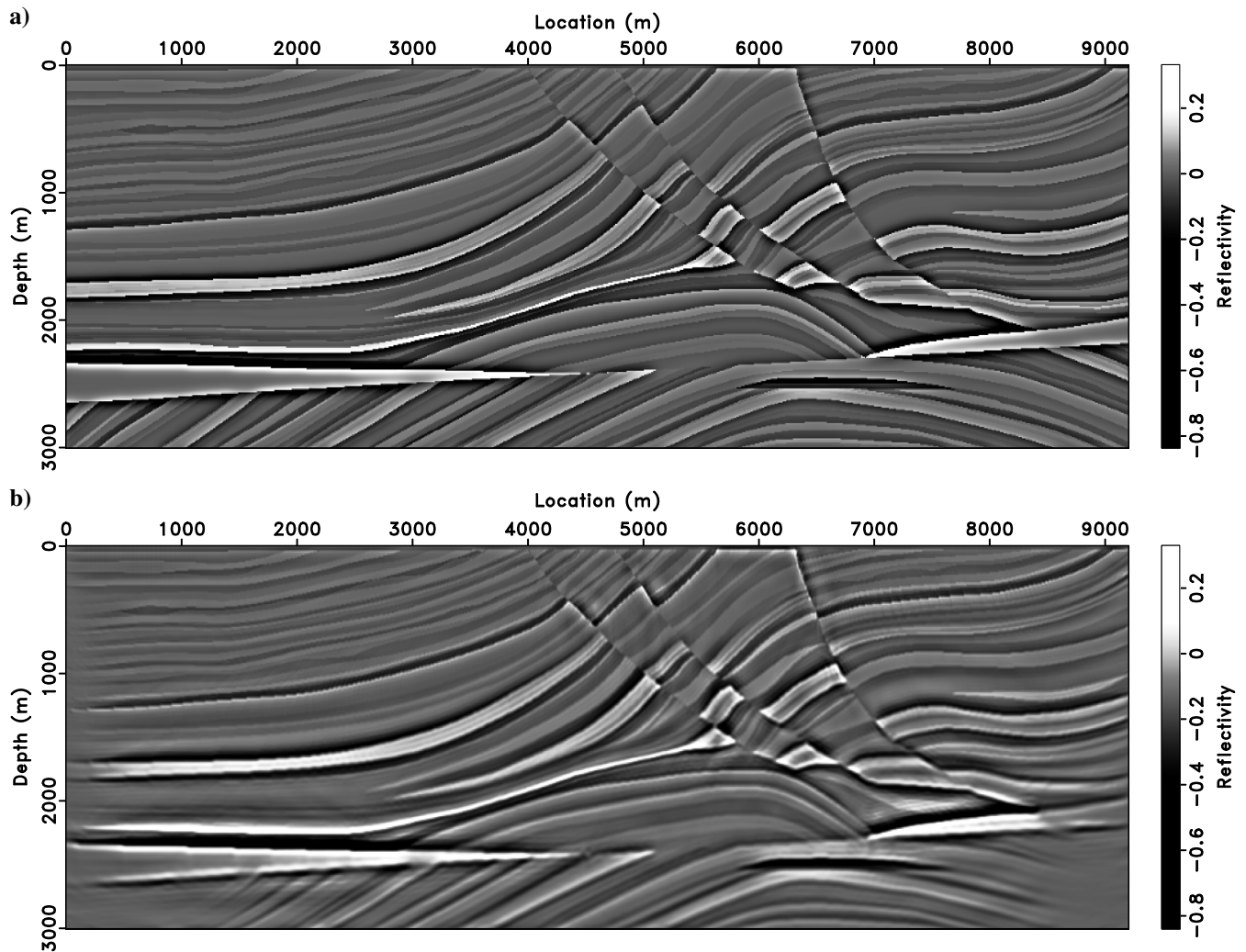


Figure 14. (a) Reflectivity model. (b) Nonextended inverted reflectivity model.



approximate inverse to the extended Born modeling operator, at least for data arising from physical (nonextended Born) modeling.

We call the process defined by equation 29 *nonextended approximate inversion*, for the purposes of this section. In our examples, we use the simplest choice of weight function,  $\phi \equiv 1$ , that is, stacking along the  $h$ -axis. The full volume (Figure 6) and the middle trace comparisons in Figure 7b illustrate the precision of the nonextended approximate inversion, and the generally low-frequency nature of the error.

The simple geometric optics computation of previous section will fail in the presence of caustics (or multipathing). The second example will show that the conclusions expressed by equations 1 and 29 are still valid even in the presence of multipathing. The background velocity model for the second example contains a low-velocity Gaussian lens. A flat horizontal reflector is placed right below the lens at a depth of 2 km. This model is very similar to the one used by Nolan and Symes (1996) and Stolk and Symes (2004). The numerical implementation has the same configuration as the first example.

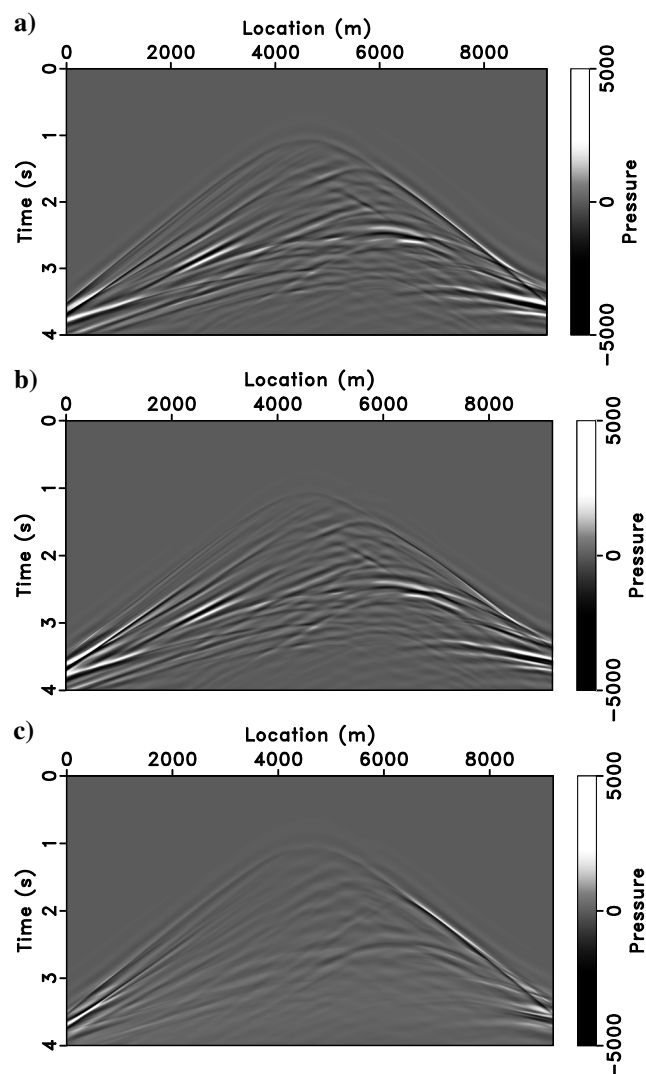


Figure 15. (a) One-shot ( $x_s = 4600$  m) Born data. (b) One-shot ( $x_s = 4600$  m) resimulated Born data. (c) Difference between the resimulated data and original data.

Because of the Gaussian lens, the rays will certainly focus and form a triplication after going through the lens. The rays and wavefronts are shown in Figure 8b. We can clearly see that this model produces multipathing and caustics.

The inverse operator defined in equation 1 produces the reflectivity model shown in Figure 9a. From the image perspective, we clearly reproduce the flat reflector below the lens with no kinematic artifacts (Stolk and Symes, 2004), consistent with kinematic prediction in Stolk et al. (2009b). Resimulation with Born modeling operator from the inverted reflectivity model predicts data very close to the input data (Figures 10a and 11a).

Finally, we apply our approximate inversion to Born data for the Marmousi model. We smooth the velocity model as the background model and take the difference as the reflectivity model, shown in Figure 12a and 12b. Born data for Marmousi model have 231 common shot gathers every 40 m, and each shot has 921 receivers every 10 m (fixed spread). The inverse operator has been applied on the Born data (middle shot shown in Figure 15a) to produce the extended inversion shown in Figure 13. The nonextended (stacked) inversion (Figure 14b) result is very similar visually to the reflectivity used in data synthesis (Figure 14a) — note that the grayscales used in these plots are identical.

The approximate inversion (Figure 14b) is only an approximation, of course. On the one hand, the inverse operator is only asymptotically correct. We can see this point from the fact that the difference between input and resimulated data is mainly of low frequency (see Figure 16b) and similarly for the model (Figure 16a). On the other hand, the theory leading to the conclusion in equation 29 explicitly ignores the possibility of scattering over  $\pi$ , that is, refraction. Some of the remaining energy in the residual data panel (Figure 15c) is refracted.

We further compare the approximate inverse operator with extended LSM (ELSM), that is, extended Born inversion via an iter-

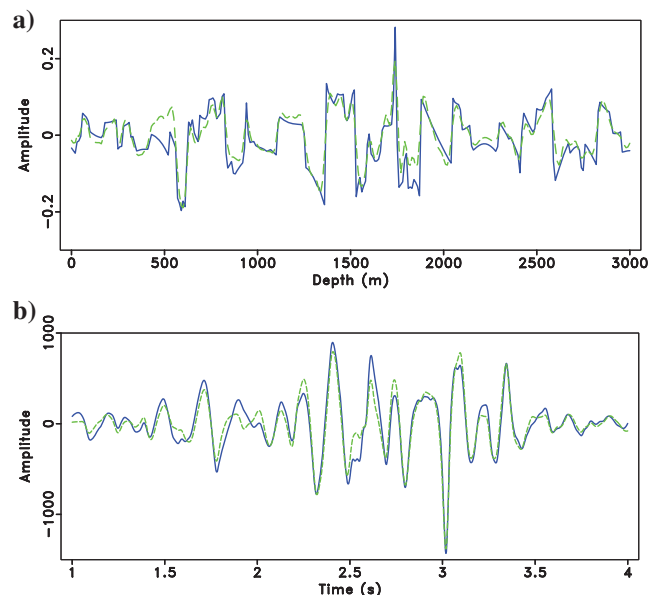


Figure 16. (a) One trace comparison ( $x = 4600$  m) between the reflectivity model (solid blue line) and nonextended inverted reflectivity model (dashed green line). (b) One trace comparison ( $x = 4600$  m) between the observed data (solid blue line) and predicted data from the inverted reflectivity model (dashed green line).

active method (we use the conjugate gradient iteration), in the same spirit as LSM (Nemeth et al., 1999; Dutta et al., 2014). A coarser grid (20 m for the spatial grid interval and 2 ms for the time interval) has been used to reduce the computational cost of ELSM. The approximate inverse operator yields the reflectivity model in Figure 17a. Starting from the approximate inversion as an initial guess, we conduct 20 iterations of ELSM (Figure 17b). On a visual comparison basis, the approximate inversion result displays no dramatic difference from ELSM result, except suffering from some low-frequency noise in the shallow part. However, a quantitative study shown in the misfit comparison (Figure 18) reveals the significant difference in terms of data misfit. The approximate inversion result produces approximately 40% relative misfit (dashed black line), whereas the 20 iteration ELSM starting from the approximate inversion result achieves roughly 10% fit error (solid black line). Note that the approximate inverse operator has almost the same computational cost as a single application of extended RTM, whereas seven or eight iterations of ELSM (each involving a migration/demigration pair) starting from zero reflectivity were required to achieve the same 40% error reduction.

As mentioned in the “Introduction” section, the weighted adjoint form of the approximate inverse operator invites inclusion in a WCG algorithm. Here, we merely present the convergence history of this algorithm, applied to the same problem. This algorithm, starting from zero reflectivity, achieves an error reduction of 10% in six or seven iterations, as opposed to the 20 equally expensive iterations of ELSM starting from the approximate inversion result, or many more iterations of ELSM starting from zero reflectivity. In 20 iterations, the WCG iteration reaches a root-mean-square error reduction factor of 3.5%. The authors will discuss this algorithm in more detail elsewhere.

## DISCUSSION

The obvious application of this construction is to accelerate iterative LSM, both extended and nonextended variants. As mentioned in the “Introduction” section, the form of the approximate inverse provides a straightforward acceleration mechanism because it shows that the extended Born modeling operator is approximately unitary in specific, computable weighted norms. The main question in this

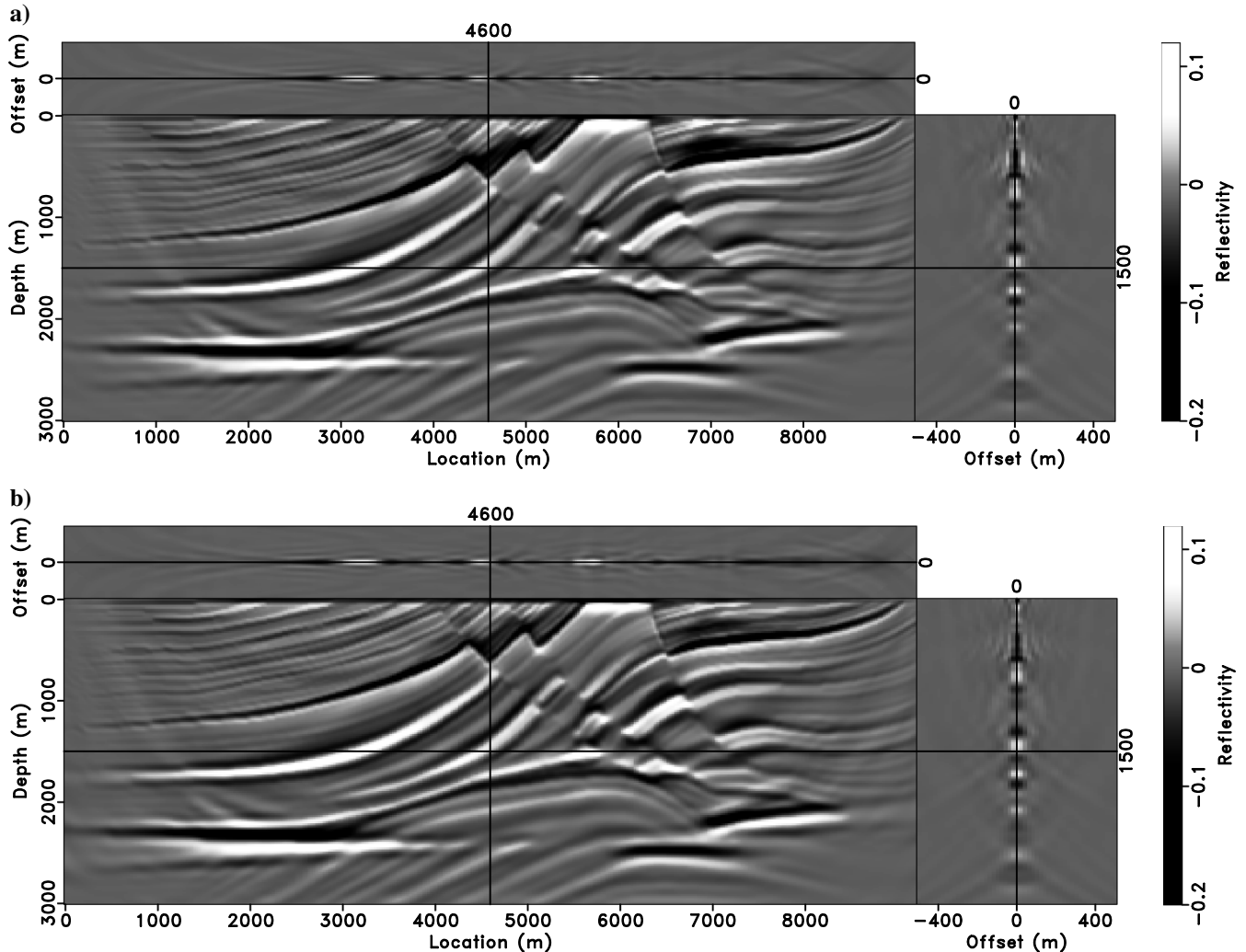


Figure 17. (a) Approximately inverted reflectivity model. (b) The 20 iterations ELSM image starting from the approximate inversion, plotted on the same grayscale.



regard is the necessary extent of the subsurface offset range. We have emphasized that this operator is negligibly more computationally intensive than *extended* RTM; however, extended RTM is considerably more expensive than ordinary RTM: The additional expense comes in the sums over offset implicit in the definition of extended modeling or migration, and it is roughly proportional to the extent of the offset axis. If an accurate velocity is available, which focuses the data, this problem is mitigated, of course. In this case, weight function  $\phi$  appearing in equation 29 is, in principle, constrained only by the requirement that  $\phi(0) = 1$ , except for finite-frequency effects. Because a weighted stack over the subsurface offset is equivalent to a weighted stack over the scattering angle, the choice of  $\phi$  will have important ramifications even for the focused case: Stacking with  $\phi = 1$  is equivalent (roughly) to using only the zero-scattering angle, or zero-offset data, and it would likely lead to suboptimal noise suppression. The design of an optimal weight  $\phi$  in the focused case, and the necessary extent of the offset range for the unfocused case and effective algorithms for determining it, are important open questions.

The operator defined in equations 1, 24, and 28 approximately inverts the extended Born modeling operator; therefore, it may be used to estimate the *extended* reflectivity in automated velocity model building methods that depend on extended inversion (Liu et al., 2013, 2014; Lameloise et al., 2015), or for amplitude versus offset studies. Note, however, that in our examples, we have adopted the approximation  $P = 1$ , which, strictly speaking, is accurate only if the image energy focuses at  $h = 0$  (that is, the velocity and data are compatible) or  $v_0$  is laterally homogeneous. In general,  $P$  is not  $\equiv 1$  in the full extended phase space volume. It is actually quite possible to remove this final impediment to asymptotic inversion:  $P$  is defined explicitly in Appendix A as a function of phase

variables, and pseudodifferential operators, such as  $W_{\text{model}}$  are effectively of low rank computable with relative efficiency, in comparison with a general matrix multiplication of the same dimension (Bao and Symes, 1996; Demanet and Ying, 2011). The improvement obtainable by one of these techniques to drop the approximation  $P = 1$  remains to be investigated. Any such computation involves the Fourier transform of the extended model or a transform of equivalent complexity. In 2D, this is already a 3D transform, whereas for three spatial dimensions, the transform is of 5D, which is a daunting prospect. However, transformation to the angle domain after Sava and Fomel (2003), considered by many authors to be an essential step (ten Kroode, 2012; Tang et al., 2013), is usually accomplished via the Fourier transform; hence, it involves the same computational load.

## CONCLUSION

A simple modification of subsurface offset extended RTM produces an asymptotic inverse to the extended Born scattering operator. Implementation of straightforward numerical experiments suggest that within its domain of applicability, this inversion operator is quite accurate. The asymptotic inverse takes the form of the adjoint of the modeling operator with respect to weighted norms in model and data spaces; hence, it offers the possibility of greatly enhancing the convergence of iterative methods for extended Born inversion.

## ACKNOWLEDGMENTS

We are grateful to the sponsors of The Rice Inversion Project (TRIP) for their long-term support and to Shell International Exploration and Production, Inc., for its support of J. Hou's Ph.D. research. We thank F. ten Kroode for inspiring our work. Thanks also go to J. Sheiman, H. Kuehl, P. Shen, and Y. Liu for very helpful discussion. We have benefited greatly from the high performance computing resources provided by the Texas Advanced Computing Center, University of Texas, Austin, and by the Rice University Research Computing Support Group. The Seismic Unix and Madagascar open-source software packages have been critically useful in our work. We appreciate the constructive comments from assistant editor M. Sacchi, associate editor I. Jones, referee J. Sheiman, and two anonymous referees.

## APPENDIX A

### PSEUDODIFFERENTIAL EXPRESSION OF THE NORMAL OPERATOR

The goal of this appendix is an expression for the slightly modified normal operator  $(I_t \bar{F}[v_0])^* (I_t \bar{F}[v])$  in the form of an oscillatory integral:

$$(I_t \bar{F}[v])^* (I_t \bar{F}[v]) \delta \bar{v}(\mathbf{x}, h) \approx \int d\mathbf{k} A(\mathbf{x}, \mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}} \hat{\delta} v(\mathbf{k}), \quad (\text{A-1})$$

over the frequency variables  $\mathbf{k} = (k_x, k_h, k_z)$ , modulo errors decaying faster than the amplitude  $A$  at large frequency. In fact, we will express the amplitude as a product of two factors: one depending on ray-trace quantities and the other depending only on values of velocity and the phase variables. The ray-dependent quantities are eventually eliminated via the identities proved in Appendix B

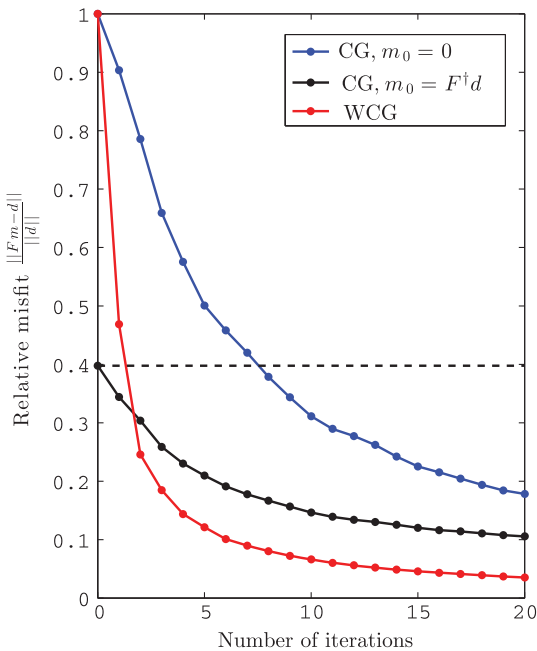


Figure 18. Relative misfit as a function of the number of iterations: The dashed black line is the approximate inversion result; the solid blue line is the ELSM result using the conjugate gradient algorithm and starting from zero reflectivity; the solid black line is the ELSM result using the conjugate gradient and starting from the approximate inversion; and the solid red line is the ELSM result using the weighted conjugate gradient (WCG) iteration.

and some further modifications of the normal operator, leading to the main result of this paper.

Begin by combining the asymptotic expressions 10 and 11, eliminating the time integral in the delta functions, and taking into account the fact that  $I_t^* = -I_t$  to obtain an integral expression for the modified normal operator introduced at the beginning of the “Theory” section:

$$(I_t \bar{F})^* (I_t \bar{F}) \delta \bar{v}(\mathbf{x}, h) \simeq -\frac{4\pi^2}{v_0(x, z)^6} \int dx_s dx_r a_s a_r \times \int d\mathbf{x}' dh' a_s' a_r' \delta(\phi(x_s, x_r, x, h, z) - \phi(x_s, x_r, x', h', z')) \delta \bar{v}(\mathbf{x}', h'), \quad (\text{A-2})$$

where  $\phi(x_s, x_r, x, h, z) = T(x_s, x - h, z) + T(x_r, x + h, z)$  is the two-way traveltime.

We recall the abbreviations

$$T_s = T(x_s, x - h, z), T_r = T(x_r, x + h, z), \quad (\text{A-3})$$

and note that

$$\frac{\partial T_s}{\partial h} = -\frac{\partial T_s}{\partial x}, \frac{\partial T_r}{\partial h} = \frac{\partial T_r}{\partial x}. \quad (\text{A-4})$$

An asymptotic evaluation of this integral follows along the lines pioneered [Beylkin \(1985\)](#), as recast by [Symes \(1998\)](#). Account for the delta function  $\delta(\phi(x_s, x_r, x, h, z) - \phi(x_s, x_r, x', h', z'))$  by writing  $z'$  as a function of  $x, h, z, x', h', x_s, x_r$  — as it will turn out, this possibility assumes that energy propagates vertically and reflectors are subhorizontal; otherwise, other space variables should be treated as dependent — then, rewrite using the inverse Fourier transform:

$$(I_t \bar{F})^* (I_t \bar{F}) \delta \bar{v}(\mathbf{x}, h) \simeq -\frac{4\pi^2}{v_0(x, z)^6} \int dx_s dx_r a_s a_r \times \int d\mathbf{x}' dh' a_s' a_r' \delta(z' - Z(x_s, x_r, x, h, z, x', h')) \times \left| \frac{\partial \phi}{\partial z'} \right|^{-1} \frac{1}{8\pi^3} \int dk_{x'} dk_{z'} dk_{h'} \hat{\delta} v(k_{x'}, k_{h'}, k_{z'}) e^{i(k_{x'} x' + k_{h'} h' + k_{z'} z')}. \quad (\text{A-5})$$

The principle of stationary phase ([Guillemin and Sternberg, 1979](#); [Bleistein et al., 2001](#)) is used to evaluate the multiple integral for a large wavenumber. This result approximates the integral of a rapidly fluctuating function  $g(\mathbf{y}) e^{i\omega \psi(\mathbf{y})}$  for large  $\omega$  by a sum of terms, one for each stationary phase point  $\mathbf{y}^*$  (that is,  $\nabla \psi(\mathbf{y}^*) = 0$ ). The general form can be written as

$$\int_{\mathbf{R}^m} d\mathbf{y} g(\mathbf{y}) e^{i\omega \psi(\mathbf{y})} \approx \sum_{\nabla \psi(\mathbf{y}^*)=0} \left( \frac{2\pi}{\omega} \right)^{\frac{m}{2}} e^{i\frac{\pi}{4} \text{sgn Hess} \psi(\mathbf{y}^*)} |\det \text{Hess} \psi(\mathbf{y}^*)|^{-\frac{1}{2}} g(\mathbf{y}^*) e^{i\omega \psi(\mathbf{y}^*)}. \quad (\text{A-6})$$

Using this approximation, the right side of equation A-5 is

$$\int dk_{x'} dk_{z'} dk_{h'} \hat{\delta} v(k_{x'}, k_{h'}, k_{z'}) \int_{\mathbf{R}^m} d\mathbf{y} g(\mathbf{y}) e^{ik_{z'} \psi(\mathbf{y})}, \quad (\text{A-7})$$

where in the general statement of the stationary phase principle A-6, we have set

$$m = 4, \mathbf{y} = (x_s, x_r, x', h'),$$

$$g(\mathbf{y}) = -\frac{1}{2\pi v_0(x, z)^6} a_s a_r a_s' a_r' \left| \frac{\partial \phi}{\partial z'} \right|^{-1},$$

$$\psi(\mathbf{y}) = \frac{k_{x'}}{k_{z'}} x' + \frac{k_{h'}}{k_{z'}} h' + Z(x_s, x_r, x, h, z, x', h'), \quad (\text{A-8})$$

and  $k_{z'}$  plays the role of large parameter  $\omega$ . To use this approximation, the Hessian

$$\text{Hess} \psi(\mathbf{y}^*) = \left( \frac{\partial^2 \psi}{\partial y_i \partial y_j} \right)_{i,j=1}^m, \quad (\text{A-9})$$

must be nonsingular at each stationary phase point and we must compute the signature and determinant of the Hessian at each such point. After standard simplifications, the stationary phase conditions are

$$\mathbf{x} = \mathbf{x}', h = h',$$

$$(k_{x'}, k_{z'}) \text{ is parallel to } \nabla_{(x', z')} \phi,$$

$$(k_{h'}, k_{z'}) \text{ is parallel to } \nabla_{(h', z')} \phi. \quad (\text{A-10})$$

During the calculation of the Hessian, the integrations are naturally paired as  $(x, x_r)$  and  $(h, x_s)$ . Each pair of integrals gives rise to a Hessian determinant factor.

The Hessian of the phase  $\psi$  can be expressed as

$$\text{Hess} = \begin{pmatrix} \frac{\partial^2 Z}{\partial x_r^2} & \frac{\partial^2 Z}{\partial x_r \partial x_s} & \frac{\partial^2 Z}{\partial x_r \partial x'} & \frac{\partial^2 Z}{\partial x_r \partial h'} \\ \frac{\partial^2 Z}{\partial x_s \partial x_r} & \frac{\partial^2 Z}{\partial x_s^2} & \frac{\partial^2 Z}{\partial x_s \partial x'} & \frac{\partial^2 Z}{\partial x_s \partial h'} \\ \frac{\partial^2 Z}{\partial x' \partial x_r} & \frac{\partial^2 Z}{\partial x' \partial x_s} & \frac{\partial^2 Z}{\partial x'^2} & \frac{\partial^2 Z}{\partial x' \partial h'} \\ \frac{\partial^2 Z}{\partial h' \partial x_r} & \frac{\partial^2 Z}{\partial h' \partial x_s} & \frac{\partial^2 Z}{\partial h' \partial x'} & \frac{\partial^2 Z}{\partial h'^2} \end{pmatrix}. \quad (\text{A-11})$$

In computing the Hessian, all derivatives must be performed before the stationary phase identities (equation A-10) are applied. The solution  $z' = Z(x_s, x_r, x, h, z, x', h')$  must satisfy the two-way travel-time equation:

$$\phi(x_s, x_r, x', h', z') = \phi(x_s, x_r, x, h, z). \quad (\text{A-12})$$

So we differentiate this equation, regarding  $x'$  as independent of  $x$ , etc., and afterwards combine with the stationary phase condition (equation A-10) to obtain

$$\frac{\partial^2 Z}{\partial x_r^2} = 0, \quad \frac{\partial^2 Z}{\partial x_s^2} = 0, \quad \frac{\partial^2 Z}{\partial x_r \partial x_s} = 0. \quad (\text{A-13})$$

Also

$$\begin{aligned}
\frac{\partial^2 \phi}{\partial x_r \partial x'} + \frac{\partial^2 \phi}{\partial x_r \partial z'} \frac{\partial Z}{\partial x'} + \frac{\partial \phi}{\partial z'} \frac{\partial^2 Z}{\partial x_r \partial x'} &= 0, \\
\frac{\partial^2 \phi}{\partial x_r \partial h'} + \frac{\partial^2 \phi}{\partial x_r \partial z'} \frac{\partial Z}{\partial h'} + \frac{\partial \phi}{\partial z'} \frac{\partial^2 Z}{\partial x_r \partial h'} &= 0, \\
\frac{\partial^2 \phi}{\partial x_s \partial x'} + \frac{\partial^2 \phi}{\partial x_s \partial z'} \frac{\partial Z}{\partial x'} + \frac{\partial \phi}{\partial z'} \frac{\partial^2 Z}{\partial x_s \partial x'} &= 0, \\
\frac{\partial^2 \phi}{\partial x_s \partial h'} + \frac{\partial^2 \phi}{\partial x_s \partial z'} \frac{\partial Z}{\partial h'} + \frac{\partial \phi}{\partial z'} \frac{\partial^2 Z}{\partial x_s \partial h'} &= 0.
\end{aligned} \quad (\text{A-14})$$

Note that the upper left  $2 \times 2$  block of the Hessian consists of zeros, and of course the Hessian is symmetric. That is, the Hessian has the following block structure:

$$\text{Hess} \sim \begin{pmatrix} 0 & A \\ A^T & B \end{pmatrix}. \quad (\text{A-15})$$

This special structure implies that the Hessian at the stationary point has exactly the same number of positive as negative eigenvalues; that is, the signature of the Hessian is zero. For details of this argument, see Symes (1998).

The block structure (equation A-15) also allows us to reduce the determinant to that of a  $2 \times 2$  matrix, squared:

$$\det \text{Hess} = - \left| \begin{array}{cc} \frac{\partial^2 Z}{\partial x' \partial x_r} & \frac{\partial^2 Z}{\partial x' \partial x_s} \\ \frac{\partial^2 Z}{\partial h' \partial x_r} & \frac{\partial^2 Z}{\partial h' \partial x_s} \end{array} \right|^2. \quad (\text{A-16})$$

The four elements of this matrix have a similar structure and can be analyzed in the same way. Take the first element as an example. Substituting equation A-14 into the first element, we get

$$\begin{aligned}
\frac{\partial^2 Z}{\partial x' \partial x_r} &= - \left( \frac{\partial \phi}{\partial z'} \right)^{-2} \left( \frac{\partial^2 \phi}{\partial x_r \partial x'} \frac{\partial \phi}{\partial z'} - \frac{\partial^2 \phi}{\partial x_r \partial z'} \frac{\partial \phi}{\partial x'} \right), \\
&= - \left( \frac{\partial \phi}{\partial z'} \right)^{-2} \det \begin{pmatrix} \frac{\partial}{\partial x_r} \nabla_{\mathbf{x}'} \phi \\ \nabla_{\mathbf{x}'} \phi \end{pmatrix}.
\end{aligned} \quad (\text{A-17})$$

Applying the same analysis on other elements and using the stationary phase conditions  $x' = x$ ,  $h' = h$ ,  $z' = z$  (because all derivatives have been computed) lead to

$$\begin{aligned}
|\det \text{Hess}|^{-1/2} &= \left[ \left( \frac{\partial \phi}{\partial z} \right)^{-4} \begin{pmatrix} \det \left( \begin{array}{c} \frac{\partial}{\partial x_r} \nabla_{(x,z)} \phi \\ \nabla_{(x,z)} \phi \end{array} \right) & \det \left( \begin{array}{c} \frac{\partial}{\partial x_s} \nabla_{(x,z)} \phi \\ \nabla_{(x,z)} \phi \end{array} \right) \\ \det \left( \begin{array}{c} \frac{\partial}{\partial x_r} \nabla_{(h,z)} \phi \\ \nabla_{(h,z)} \phi \end{array} \right) & \det \left( \begin{array}{c} \frac{\partial}{\partial x_s} \nabla_{(h,z)} \phi \\ \nabla_{(h,z)} \phi \end{array} \right) \end{pmatrix} \right]^{-1}.
\end{aligned} \quad (\text{A-18})$$

We write  $s = 1/v_0$  for slowness, and  $s_{\pm} = s(x \pm h, z)$ . Note that the eikonal equation asserts that

$$\nabla_{x,z} T_s \cdot \nabla_{x,z} T_s = s_-^2, \quad \nabla_{x,z} T_r \cdot \nabla_{x,z} T_r = s_+^2. \quad (\text{A-19})$$

Because the lengths of the traveltimes gradients are independent of the source and receiver coordinates, their derivatives with respect to these coordinates must be orthogonal to the gradients. A simple way to express this fact is to write

$$\nabla_{x,z} T_s = s_- (\sin \alpha_s, \cos \alpha_s), \quad (\text{A-20})$$

$$\nabla_{x,z} T_r = s_+ (\sin \alpha_r, \cos \alpha_r), \quad (\text{A-21})$$

so,

$$\begin{aligned}
\frac{\partial}{\partial x_s} \frac{\partial T_s}{\partial x} &= \frac{\partial T_s}{\partial z} \frac{\partial \alpha_s}{\partial x_s}, & \frac{\partial}{\partial x_s} \frac{\partial T_s}{\partial z} &= - \frac{\partial T_s}{\partial x} \frac{\partial \alpha_s}{\partial x_s}, \\
\frac{\partial}{\partial x_r} \frac{\partial T_r}{\partial x} &= \frac{\partial T_r}{\partial z} \frac{\partial \alpha_r}{\partial x_r}, & \frac{\partial}{\partial x_r} \frac{\partial T_r}{\partial z} &= - \frac{\partial T_r}{\partial x} \frac{\partial \alpha_r}{\partial x_r}.
\end{aligned} \quad (\text{A-22})$$

The elements of the matrix in equation A-18 simplify due to A-22: For example, the (1,1) element becomes

$$\begin{aligned}
\left| \begin{array}{cc} \frac{\partial}{\partial x_r} \frac{\partial T_r}{\partial x} & \frac{\partial}{\partial x_r} \frac{\partial T_r}{\partial z} \\ \frac{\partial T_r}{\partial x} + \frac{\partial T_s}{\partial x} & \frac{\partial T_r}{\partial z} + \frac{\partial T_s}{\partial z} \end{array} \right| &= \frac{\partial \alpha_r}{\partial x_r} \left| \begin{array}{cc} \frac{\partial T_r}{\partial z} & - \frac{\partial T_r}{\partial x} \\ \frac{\partial T_r}{\partial x} + \frac{\partial T_s}{\partial x} & \frac{\partial T_r}{\partial z} + \frac{\partial T_s}{\partial z} \end{array} \right| \\
&= s_-^2 + \nabla_{x,z} T_r \cdot \nabla_{x,z} T_s,
\end{aligned} \quad (\text{A-23})$$

by virtue of the eikonal equation A-19. Evaluating the other elements similarly

$$\begin{aligned}
|\det \text{Hess}|^{-1/2} &= \left( \frac{\partial \phi}{\partial z} \right)^4 \\
&\times \left[ \frac{\partial \alpha_s \partial \alpha_r}{\partial x_s \partial x_r} \left( s_+^2 + \nabla_{x,z} T_r \cdot \nabla_{x,z} T_s \right) \left( s_-^2 + \nabla_{x,z} T_r \cdot \nabla_{x,z} T_s \right) - s_-^2 - \nabla_{x,z} T_r \cdot \nabla_{h,z} T_s \right]^{-1}, \\
&= - \frac{1}{2} \left( \frac{\partial \phi}{\partial z} \right)^4 \left( \frac{\partial \alpha_s \partial \alpha_r}{\partial x_s \partial x_r} \right)^{-1} \\
&\times [(s_-^2 s_+^2 + (\nabla_{x,z} T_r \cdot \nabla_{x,z} T_s)(\nabla_{x,z} T_r \cdot \nabla_{h,z} T_s) \\
&+ (s_-^2 + s_+^2)(\nabla_{x,z} T_r \cdot \nabla_{x,z} T_s + \nabla_{x,z} T_r \cdot \nabla_{h,z} T_s))]^{-1}, \\
&= - \frac{1}{2} \left( \frac{\partial \phi}{\partial z} \right)^4 \left( \frac{\partial \alpha_s \partial \alpha_r}{\partial x_s \partial x_r} \right)^{-1} \\
&\times \left[ \left( s_-^2 \left( \frac{\partial T_r}{\partial z} \right)^2 + s_+^2 \left( \frac{\partial T_s}{\partial z} \right)^2 \right) + (s_-^2 + s_+^2) \left( \frac{\partial T_s}{\partial z} \frac{\partial T_r}{\partial z} \right) \right]^{-1}.
\end{aligned} \quad (\text{A-24})$$

Apart from the angle derivatives, this expression is actually algebraic in the phase variables and  $s_{\pm}$ . To see this, we invoke the remaining stationary phase conditions (equation A-10) pertaining to the phase variables. Adding and subtracting ratios equivalent to these conditions, we obtain

$$\begin{aligned}
\frac{\partial T_r}{\partial x} &= \frac{1}{2} \frac{k_x + k_h}{k_z} \frac{\partial \phi}{\partial z}, \\
\frac{\partial T_s}{\partial x} &= \frac{1}{2} \frac{k_x - k_h}{k_z} \frac{\partial \phi}{\partial z}.
\end{aligned} \quad (\text{A-25})$$

Then, the eikonal equation implies that

$$\begin{aligned} s_+^2 - \frac{1}{4} \left[ \frac{k_x + k_h}{k_z} \right]^2 \left( \frac{\partial \phi}{\partial z} \right)^2 &= \left( \frac{\partial T_r}{\partial z} \right)^2, \\ s_-^2 - \frac{1}{4} \left[ \frac{k_x - k_h}{k_z} \right]^2 \left( \frac{\partial \phi}{\partial z} \right)^2 &= \left( \frac{\partial T_s}{\partial z} \right)^2. \end{aligned} \quad (\text{A-26})$$

For convenience, we set

$$a_{\pm} = \frac{1}{4} \left[ \frac{k_x \pm k_h}{k_z} \right]^2, \quad \zeta_{\pm} = \frac{\partial T_r}{\partial z} \pm \frac{\partial T_s}{\partial z}, \quad (\text{A-27})$$

so,  $\zeta_+ = \partial \phi / \partial z$ .

Adding the two equations A-25, and rearranging, we obtain  $\zeta_+^2$  in terms of  $\zeta_-^2$ :

$$\zeta_-^2 = 2(s_+^2 + s_-^2) - [2(a_+ + a_-) + 1]\zeta_+^2. \quad (\text{A-28})$$

Subtracting the two equations A-25, squaring the result, eliminating  $\zeta_-^2$  using equation A-26, and rearranging yields a quadratic equation for  $\zeta_+^2$ :

$$a(\zeta_+^2)^2 + b\zeta_+^2 + c = 0, \quad (\text{A-29})$$

in which

$$a = (a_+ - a_-)^2 + 2(a_+ + a_-) + 1 = \frac{k_{xz}^2 k_{hz}^2}{k_z^4}, \quad (\text{A-30})$$

$$b = -2 \left[ (s_+^2 - s_-^2) \frac{k_x k_h}{k_z^2} + (s_+^2 + s_-^2) \right], \quad c = (s_+^2 - s_-^2)^2. \quad (\text{A-31})$$

Because we have assumed from the beginning that  $\partial \phi / \partial z > 0$  in the region of interest, also when  $h = 0$ , so that  $s_+ = s_-$ , the choice of root is fixed

$$\left( \frac{\partial \phi}{\partial z} \right)^2 = \zeta_+^2 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}. \quad (\text{A-32})$$

**Remark:** A very similar argument occurs in [ten Kroode \(2012\)](#), leading up to equation A-14.

We can now assemble the amplitude in the stationary phase approximation (equation A-6), using the integrand given in equation A-8, the expression for the Hessian, and the frequency factor: We obtain for the integrand in equation A-6:

$$\begin{aligned} \frac{2\pi s^6}{k_z^2} a_r^2 a_s^2 \left( \frac{\partial \phi}{\partial z} \right)^{-1} |\det \text{Hess}|^{-1/2} \\ = \frac{a_r^2 a_s^2}{k_z^2} \left( \frac{\partial \alpha_s}{\partial x_s} \frac{\partial \alpha_r}{\partial x_r} \right)^{-1} \tilde{P}(x, h, z, k_x, k_h, k_z), \end{aligned} \quad (\text{A-33})$$

in which

$$\begin{aligned} \tilde{P} = -\pi s^4 \left( \frac{\partial \phi}{\partial z} \right)^3 \left[ \left( \left( \frac{s_-}{s} \right)^2 \left( \frac{\partial T_r}{\partial z} \right)^2 + \left( \frac{s_+}{s} \right)^2 \left( \frac{\partial T_s}{\partial z} \right)^2 \right) \right. \\ \left. + \left( \left( \frac{s_-}{s} \right)^2 + \left( \frac{s_+}{s} \right)^2 \right) \left( \frac{\partial T_s}{\partial z} \frac{\partial T_r}{\partial z} \right) \right]^{-1}. \end{aligned} \quad (\text{A-34})$$

From the defining relations A-23, A-25, A-29, A-30, and A-31, it follows that  $\tilde{P}$  is homogeneous of order zero in the phase variables  $k_x, k_h, k_z$ . The various components of  $\tilde{P}$  simplify considerably when  $h = 0$  (important because physical reflectivities are supported there). The term in brackets in equation A-34 becomes precisely  $(\partial \phi / \partial z)^2$ , so

$$\tilde{P}(x, z, 0, k_x, k_h, k_z) = -\pi s^4 \frac{\partial \phi}{\partial z} = -2\pi s^5 \frac{k_z^2}{k_{xz} k_{hz}}. \quad (\text{A-35})$$

Dividing  $\tilde{P}$  by the right side in equation A-35 produces  $P$ , also homogeneous of order zero in the frequency variables and  $\equiv 1$  for  $h = 0$ . Thus, we arrive at the integral representation (equation 12) in the ‘‘Theory’’ section, with amplitude

$$\frac{2\pi s^6}{k_z^2} a_r^2 a_s^2 \left( \frac{\partial \phi}{\partial z} \right)^{-1} |\det \text{Hess}|^{-1/2} = -\frac{2\pi s^5}{k_{xz} k_{hz}} P a_r^2 a_s^2 \left( \frac{\partial \alpha_s}{\partial x_s} \frac{\partial \alpha_r}{\partial x_r} \right)^{-1}. \quad (\text{A-36})$$

Note that all factors on the right side are functions of the phase variables, although some of them are apparently to be determined by ray tracing, as was claimed at the beginning of this appendix.

## APPENDIX B

### ANALYSIS OF GEOMETRIC AMPLITUDES

As far as we know, the relation explained in this appendix appeared first in [Zhang et al. \(2005\)](#). We rederive the relation in this appendix from a different perspective for the convenience of the reader.

The derivation starts from the transport equation. The transport equation in divergence form is

$$\nabla \cdot (a^2 \nabla \tau) = 0. \quad (\text{B-1})$$

Consider a region  $R$  formed by two rays radiating from the same point (see Figure B-1). Truncate this region with two lines  $l$  and  $L$

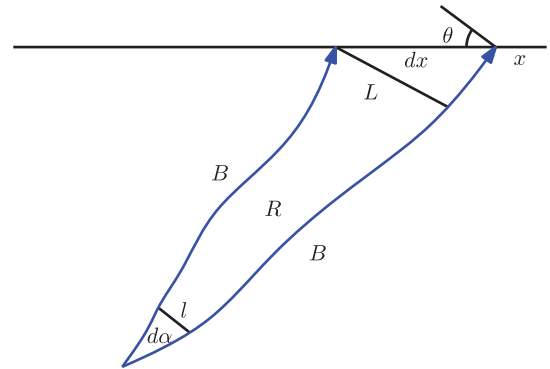


Figure B-1. Sketch of geometric amplitudes analysis.

normal to the rays. Denoted by  $\alpha$ , the angle subtended between the first ray and the vertical.

Apply the 2D divergence theorem in the enclosed region  $R$ . We can get

$$\iint_R \nabla \cdot (a^2 \nabla \tau) dR = \oint_C a^2 \nabla \tau \cdot \mathbf{n} dC = 0, \quad (\text{B-2})$$

where  $\mathbf{n}$  is the outward normal vector to the boundary  $C = B \cup l \cup L$

- 1) For  $\mathbf{x}$  on  $B$ , the normal vector is perpendicular to the ray,  $\nabla \tau \cdot \mathbf{n} = 0$ .
- 2) For  $\mathbf{x}$  on  $l$  and  $L$ ,  $\mathbf{n}$  is parallel to the ray,  $\nabla \tau \cdot \mathbf{n} = \pm |\nabla \tau|$ .

Therefore,

$$\int_L A^2 |\nabla \tau| dL = \int_l a^2 |\nabla \tau| dl, \quad (\text{B-3})$$

which immediately leads to

$$A^2 = a^2 \frac{V}{v} \frac{dl}{dL} = a^2 \frac{V}{v} \frac{dl}{d\alpha} \frac{1}{\cos \theta} \frac{d\alpha}{dx}. \quad (\text{B-4})$$

We have denoted the velocity and amplitude near where the ray starts as  $v, a$  and those near where the ray terminates as  $V, A$ . In the vicinity of the starting point, the 2D geometric amplitude has the constant-velocity approximation

$$a^2 \approx \frac{v}{8\pi^2 r} = \frac{v}{8\pi^2} \frac{d\alpha}{dl}. \quad (\text{B-5})$$

So, in particular for the receiver ray,

$$a_r^2 = \frac{1}{8\pi^2} \frac{v_r}{\cos \theta_r} \frac{d\alpha_r}{dx_r}, \quad (\text{B-6})$$

and similarly for the source ray,

$$a_s^2 = \frac{1}{8\pi^2} \frac{v_s}{\cos \theta_s} \frac{d\alpha_s}{dx_s}. \quad (\text{B-7})$$

## REFERENCES

- Bao, G., and W. Symes, 1996, Computation of pseudo-differential operators: SIAM Journal on Scientific Computation, **17**, 416–429.
- Baysal, E., D. D. Kosloff, and J. W. C. Sherwood, 1983, Reverse time migration: Geophysics, **48**, 1514–1524, doi: [10.1190/1.1441434](https://doi.org/10.1190/1.1441434).
- Bednar, J. B., and C. J. Bednar, 2006, Two-way vs. One-way: A case study style comparison: 76th Annual International Meeting, SEG, Expanded Abstracts, 2343–2347.
- Beylkin, G., 1985, Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized radon transform: Journal of Mathematical Physics, **26**, 99–108, doi: [10.1063/1.526755](https://doi.org/10.1063/1.526755).
- Biondi, B., and A. Almomin, 2014, Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion: Geophysics, **79**, no. 3, WA129–WA140, doi: [10.1190/geo2013-0340.1](https://doi.org/10.1190/geo2013-0340.1).
- Bleistein, N., 1987, On the imaging of reflectors in the earth: Geophysics, **52**, 931–942, doi: [10.1190/1.1442363](https://doi.org/10.1190/1.1442363).
- Bleistein, N., J. Cohen, and J. Stockwell, 2001, Mathematics of multidimensional seismic imaging, migration, and inversion: Springer.
- Bleistein, N., Y. Zhang, S. Xu, G. Zhang, and S. H. Gray, 2005, Migration/inversion: Think image point coordinates, process in acquisition surface coordinates: Inverse Problems, **21**, 1715–1744, doi: [10.1088/0266-5611/21/5/013](https://doi.org/10.1088/0266-5611/21/5/013).
- Bourgeois, A., B. Jiang, and P. Lailly, 1989, Linearized inversion: A significant step beyond pre-stack migration: Geophysical Journal International, **99**, 435–445, doi: [10.1111/j.1365-246X.1989.tb01700.x](https://doi.org/10.1111/j.1365-246X.1989.tb01700.x).
- Clairbout, J. F., 1985, Imaging the earth's interior: Blackwell Scientific Publishers.
- Courant, R., and D. Hilbert, 1962, Methods of mathematical physics: Wiley-Interscience.
- De Hoop, M. V., and N. Bleistein, 1997, Generalized Radon transform inversions for reflectivity in anisotropic elastic media: Inverse Problems, **16**, 669–690, doi: [10.1088/0266-5611/16/3/009](https://doi.org/10.1088/0266-5611/16/3/009).
- Demaret, L., and L. Ying, 2011, Discrete symbol calculus: SIAM Review, **53**, 71–104, doi: [10.1137/080731311](https://doi.org/10.1137/080731311).
- Dutta, G., Y. Huang, W. Dai, X. Wang, and G. T. Schuster, 2014, Making the most out of the least squares migration: 84th Annual International Meeting, SEG, Expanded Abstracts, 4405–4410.
- Gelfand, I., and G. Shilov, 1958, Generalized functions: Academic Press.
- Guillemin, V., and S. Sternberg, 1979, Geometric asymptotics: American Mathematical Society.
- Hou, J., and W. Symes, 2014, An approximate inverse to the extended Born modeling operator: 84th Annual International Meeting, SEG, Expanded Abstracts, 3784–3789.
- Kern, M., and W. Symes, 1994, Inversion of reflection seismograms by differential semblance analysis: Algorithm structure and synthetic examples: Geophysical Prospecting, **99**, 565–614, doi: [10.1111/j.1365-2478.1994.tb00231.x](https://doi.org/10.1111/j.1365-2478.1994.tb00231.x).
- Lameloise, C.-A., H. Chauris, and M. Noble, 2015, Improving the gradient of the image-domain objective function using quantitative migration for a more robust migration velocity analysis: Geophysical Prospecting, **63**, 391–404, doi: [10.1111/1365-2478.12195](https://doi.org/10.1111/1365-2478.12195).
- Liu, Y., W. Symes, and Z. Li, 2014, Inversion velocity analysis via differential semblance optimization: 76th Annual International Meeting, EAGE, Extended Abstracts, doi: [10.3997/2214-4609.20141578](https://doi.org/10.3997/2214-4609.20141578).
- Liu, Y., W. W. Symes, Y. Huang, and Z. Li, 2013, Linearized extended waveform inversion via differential semblance optimization in the depth oriented extension: 83rd Annual International Meeting, SEG, Expanded Abstracts, 4869–4874.
- Loewenthal, D., and I. Mufti, 1983, Reversed time migration in spatial frequency domain: Geophysics, **48**, 627–635, doi: [10.1190/1.1441493](https://doi.org/10.1190/1.1441493).
- Mulder, W. A., and R.-E. Plessix, 2004, A comparison between one-way and two-way wave equation migration: Geophysics, **69**, 1491–1504, doi: [10.1190/1.1836822](https://doi.org/10.1190/1.1836822).
- Nemeth, T., C. Wu, and G. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, 208–221, doi: [10.1190/1.1444517](https://doi.org/10.1190/1.1444517).
- Nolan, C. J., and W. Symes, 1996, Imaging and coherency in complex structure: 66th Annual International Meeting, SEG, Expanded Abstracts, 359–363.
- Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503, doi: [10.1111/j.1365-246X.2006.02978.x](https://doi.org/10.1111/j.1365-246X.2006.02978.x).
- Rickett, J., and P. Sava, 2002, Offset and angle-domain common image-point gathers for shot profile migration: Geophysics, **67**, 883–889, doi: [10.1190/1.1484531](https://doi.org/10.1190/1.1484531).
- Sava, P., and S. Fomel, 2003, Angle-domain common-image gathers by wavefield continuation methods: Geophysics, **68**, 1065–1074, doi: [10.1190/1.1581078](https://doi.org/10.1190/1.1581078).
- Stolk, C. C., M. V. de Hoop, and T. P. M. Op't Root, 2009a, Reverse time migration inversion from single-shot data: 79th Annual International Meeting, SEG, Expanded Abstracts, 2995–2999.
- Stolk, C. C., M. V. de Hoop, and W. Symes, 2009b, Kinematics of shot-geophone migration: Geophysics, **74**, no. 6, WCA19–WCA34, doi: [10.1190/1.3256285](https://doi.org/10.1190/1.3256285).
- Stolk, C. C., and W. Symes, 2004, Kinematic artifacts in prestack depth migration: Geophysics, **69**, 562–575, doi: [10.1190/1.1707076](https://doi.org/10.1190/1.1707076).
- Sun, D., and W. Symes, 2012, Waveform inversion via nonlinear differential semblance optimization: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: [10.1190/segam2012-1190.1](https://doi.org/10.1190/segam2012-1190.1).
- Symes, W., 1998, Mathematics of reflection seismology, <http://www.trip.caam.rice.edu>, accessed 15 February 2014.
- Symes, W., 2008, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **56**, 765–790, doi: [10.1111/j.1365-2478.2008.00698.x](https://doi.org/10.1111/j.1365-2478.2008.00698.x).
- Symes, W., 2009, The seismic reflection inverse problem: Inverse Problems, **25**, 123008, doi: [10.1088/0266-5611/25/12/123008](https://doi.org/10.1088/0266-5611/25/12/123008).
- Tang, B., S. Xu, and Y. Zhang, 2013, 3D angle gathers with plane-wave reverse time migration: Geophysics, **78**, no. 2, S117–S123, doi: [10.1190/geo2012-0313.1](https://doi.org/10.1190/geo2012-0313.1).
- Taylor, M., 1981, Pseudodifferential operators: Princeton University Press.
- Ten Kroode, F., 2012, A wave-equation-based Kirchhoff operator: Inverse Problems, **28**, 115013, doi: [10.1088/0266-5611/28/1/115013](https://doi.org/10.1088/0266-5611/28/1/115013).



- Whitmore, N. D., 1983, Iterative depth migration by backward time propagation: 53rd Annual International Meeting, SEG, Expanded Abstracts, 382–385.
- Xu, S., Y. Zhang, and B. Tang, 2011, 3D angle gathers from reverse time migration: *Geophysics*, **76**, no. 2, S77–S92, doi: [10.1190/1.3536527](https://doi.org/10.1190/1.3536527).
- Zhang, Y., and J. Sun, 2008, Practical issues of reverse time migration-true-amplitude gathers, noise removal and harmonic-source encoding: Presented at 76th Annual International Meeting, EAGE, Extended Abstracts, doi: [10.3997/2214-4609.20147708](https://doi.org/10.3997/2214-4609.20147708).
- Zhang, Y., J. Sun, and S. H. Gray, 2007, Reverse time migration: Amplitude and implementation issues: 77th Annual International Meeting, SEG, Expanded Abstracts, 2145–2149.
- Zhang, Y., G. Zhang, and N. Bleistein, 2005, Theory of true-amplitude one-way wave equations and true-amplitude common-shot migration: *Geophysics*, **70**, no. 4, E1–E10, doi: [10.1190/1.1988182](https://doi.org/10.1190/1.1988182).